THE DYNAMICS AND EVOLUTION OF SOCIAL SYSTEMS

New Foundations of a Mathematical Sociology

by

JÜRGEN KLÜVER
University of Essen, Germany

SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

THEORY AND DECISION LIBRARY

General Editors: W. Leinfellner (*Vienna*) and G. Eberlein (*Munich*)

Series A: Philosophy and Methodology of the Social Sciences

Series B: Mathematical and Statistical Methods

Series C: Game Theory, Mathematical Programming and Operations Research

SERIES A: PHILOSOPHY AND METHODOLOGY OF THE SOCIAL SCIENCES

VOLUME 29

Series Editor: W. Leinfellner (Technical University of Vienna), G. Eberlein (Technical University of Munich); *Editorial Board:* R. Boudon (Paris), M. Bunge (Montreal), J. S. Coleman (Chicago), J. Götschl (Graz), L. Kern (Pullach), I. Levi (New York), R. Mattessich (Vancouver), B. Munier (Cachan), J. Nida-Rümelin (Göttingen), A. Rapoport (Toronto), A. Sen (Cambridge, U.S.A.), R. Tuomela (Helsinki), A. Tversky (Stanford).

Scope: This series deals with the foundations, the general methodology and the criteria, goals and purpose of the social sciences. The emphasis in the Series A will be on well-argued, thoroughly analytical rather than advanced mathematical treatments. In this context, particular attention will be paid to game and decision theory and general philosophical topics from mathematics, psychology and economics, such as game theory, voting and welfare theory, with applications to political science, sociology, law and ethics.

The titles published in this series are listed at the end of this volume.

THE DYNAMICS AND EVOLUTION OF SOCIAL SYSTEMS

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-90-481-5514-9 ISBN 978-94-015-9570-4 (eBook)
DOI 10.1007/978-94-015-9570-4

Printed on acid-free paper

All Rights Reserved
2000 Springer Science+Business Media Dordrecht
Originally published by Kluwer Academic Publishers in 2000
Softcover reprint of the hardcover 1st edition 2000
No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the copyright owner.

PREFACE

When I started with this book several years ago I originally intended to write an introduction to mathematical systems theory for social scientists. Yet the more I thought about systems theory on the one side and theoretical sociology on the other the more I became convinced that the classical mathematical tools are not very well suited for the problems of sociology. Then I became acquainted with the researches on complex systems by the Santa Fé Institute and in particular with cellular automata, Boolean networks and genetic algorithms. These mathematically very simple but extremely efficient tools are, in my opinion, very well appropriate for modeling social dynamics. Therefore I tried to reformulate several classical problems of theoretical sociology in terms of these formal systems and outline new possibilities for a mathematical sociology which is able to join immediately on the great traditions of theoretical sociology. The result is this book; whether I succeeded with it is of course up to the readers.

As the readers will perceive, the book could not have been written by me alone but only by the joint labors of the computer group at the Interdisciplinary Center of Research in Higher Education at the University of Essen. The members of the group, Christina Stoica, Jörn Schmidt and Ralph Kier, are named in several subchapters as co-authors. Yet even more important than their contributions to this book were the permanent discussions with them and their patience with my new and very speculative ideas. Many thanks.

I am also very grateful to Helga Andresen who convinced me to read Bateson again and discussed often his ideas with me. She also gave me very valuable information about theoretical linguistics, a field of which I know only little.

The translation from the German original was done by Frank Born, whose results I can only admire. Corrections as a native speaker were done by Brandon Claycomb - thanks also to him. Any mistakes in this book are, of course, my responsibility. The main work on the manuscript was done by Irmgard Siekerkotte: she never complained and continually gave her best which was always more than enough.

Meanwhile some ideas and results of this book have been published in different journals. Several editors and reviewers acquainted me with information about articles and books I did not know and which I have included in this book. In particular I am grateful for such information to Tom Fararo, Patrick Doreian, Steve Lansing, Dwight Read and Klaus Troitzsch. Yaneer Bar-Yam invited me to NECSI - New England Complex Systems Institute - and gave me the opportunity to discuss some

of my ideas with a fascinating interdisciplinary audience. *Ut vivant, crescant, floreant.*

I dedicate the book to my son Rasmus, who is not little any more, and to the memory of my father, who died on New Years Eve 1998.

And now: *habeant sua fata libelli.*

Table of contents

1. Introduction: Systems, Theory, Computer, and Sociology **1**

1.1. Sociology and Systems Theory 4

1.2. The New Sciences of Complexity 7

1.3. Social and Cognitive Systems: the Systemic Objects of the Humanities and the Social Sciences 15

2. State, Evolution, and Complexity: Building Blocks of the Theories of Complex Systems **25**

2.1. General Concepts 25

2.2. Adaptation and Self-organization - An Eternal Golden Braid 28

2.3. Evolution, Learning, and Selfmodeling: Self - referential Dynamics 35

2.4. The Memory of Systems or: Is It Possible to Learn from History? 41

2.5. Gödel, Turing, and Münchhausen: The Paradigm of Universal Computability 50

2.6. Complexity and Emergence: Concepts and their Vagueness 57

2.7. Systemic Thinking: The Kantian Stance and Functionalism 68

3. The Dynamics and Evolution of Formal Systems **79**

3.1. Cellular Automata and Boolean Nets: The Paradigm of Self-organization 81

3.1.1. The Logic of the Chessboard: Cellular Automata (CAs) 81

3.1.2. Boolean Networks (BNs): Making (Propositional) Logic Dynamic 96

3.2. Genetic Algorithms (GAs): Self-organization Through Adaptation 106

3.3. Hybrid Systems: Dynamics and Metadynamics 122
3.3.1. Rule Variation, Environment, and Hybrid Structures 123
3.3.2. Digression: Conditions of Rule Variation or: Quis custodiet ipsos custodes? 135
3.3.3. The Parameterization of Rule Spaces: Variability Degrees and Meta-Parameters (together with Jörn Schmidt and Ralph Kier) 141
3.3.4. The Mathematics of Selfreferentiality 158

4. Building Blocks of a Mathematical Sociology 161

4.1. Systemic Stagnations, Regressions, and Conservatisms (together with Jörn Schmidt) 164
4.2. Selfreferentiality as Self-modeling (together with Christina Stoica and Jörn Schmidt) 173
4.3. Game Strategies and System Dynamics: Some Thoughts on the Relations of the Theory of Games and Systems Theory (together with Ralph Kier) 181
4.4. The Charm of a Discrete Geometry (together with Jörn Schmidt) 209
4.4.1. The Topology and Metric of Social Spaces 210
4.4.2. Spatial Dimensions and Social Differentiations 223
4.5. Some Thoughts on the Development of the New 234

5. Rules, Universals, and Questions of Research - A Conclusion That Is Not An Ending 241

5.1. The Regularities of Social Action: Some Aspects of the Social Rule Concept 245
5.2. Social Universals and (Biological) Constraints 257
5.3. Conclusion and Prospects 264

References 273

Index 285

1. INTRODUCTION: SYSTEMS, THEORY, COMPUTER, AND SOCIOLOGY

The underlying main idea of this work is, like all basic ideas, very simple and, strictly speaking, as old as theoretical sociology itself. It simply consists in the fact that people construct social reality by interacting in accordance with specific rules. By this means, contexts of action and interaction emerge that are described, in a now familiar terminology, as "systems". These systems, generated by rule-governed interactions, retroact on every situation of action, which in turn modifies every single action and interaction, and so on. For the social scientist, it is merely a matter of methodological approach, and not of social ontology, to derive one's basic assumptions from the realm of elementary actions and interactions or from the higher level of systems. In either case, the interdependence between the level of the local action and that of systems has to be taken into consideration. The decisive consideration is this: on the basis of which rules are (local) interactions performed, and what kind of system dynamics result from these rules? Hence, the task of sociology can be seen as that of analyzing the logic and co*nsequences of social rule systems*.

So far, so simple - and so classical. Practically every classical theoretical approach to the social sciences could be translated into this terminology - including the rules of class struggle (Klassenkampf), constituted by the dialectics of relations of production (Produktionsverhältnisse) and forces of production (Produktivkräfte) in Marx; the rules of rational action in Weber; the duality between communicative and strategic-instrumental action in Habermas; and the recursive autopoiesis in Luhmann's communication systems. Numerous social theorists have varied this idea, continually providing new terminologies, which certainly yielded increasingly new insights into the "complexity" of social systems.

In any case, it would be neither original nor particularly fruitful to simply add yet another approach to those already existing, which merely would operate on the basis of a more or less modified terminology with the same methods as before. The aim of this study is to follow a new methodological course. On the basis of recent developments of both general and natural scientific systems theory on the one hand, and elaborated computer programs on the other hand, it has become possible to make the above mentioned idea the starting point of formal logical and mathematical analysis. Rules as well as the systems constituted by them can be formally defined, patterns of behavior of different systems can be

J. Klüver, *The Dynamics and Evolution of Social Systems*
© Springer Science+Business Media Dordrecht 2000

studied in computer modelings, and, additionally, mathematical characteristics of rule systems - though still very provisional - can be classified. It has to be emphasized right from the start that this not only applies to systems in the natural sciences, but equally to social systems - though problems in the latter case are of course much more complicated than in case of the relatively simple systems of nature.

If this approach bears close examination - and there is a lot to be said for it at present - this would amount to a big step towards the old sociologists' dream of grasping social reality with the methodological precision of the natural sciences. What is not intended here, of course, is any kind of scientific reductionism which simply adopts terms and established methods of the natural sciences literally. On the contrary, if one analyzes social systems meticulously, it becomes apparent that classical as well as more recent concepts of natural scientific systems theory cannot be implemented directly in the social sciences but instead have to be extended considerably. The particularity of the social cannot be reduced to scientific terminology, though it provides a great deal of the terminological and methodological frame that is needed. To bring this discussion to a point: a formally strict theory of social systems that could not integrate the theoretical level of the great classic approaches to sociology might not be completely worthless, but certainly would be neither sufficient nor of much interest to the majority of social theorists. When constructing and analyzing formal models of social life, which is the main focus of this work, one has to orient oneself toward the terminological and theoretical insights that have been achieved from Marx to Luhmann. The efficiency, that is to say the sociological adequacy of formal models, will have to be measured by that standard. By no means does this entail perpetuating the traditional mistake of reductionist approaches, i.e. accepting goals as scientific only if they are attainable by narrow means, a view that requires that "scientific" problems themselves must be suited to scientific methodology. One should instead assert the opposite: the method of strict analysis must adapt to the problems of the social world. Thus, in a very general sense Althusser's famous advice still holds true for the constructors of sociologically interesting formal models: Lire le Capital.

This book certainly does not provide a complete and definite theory of social systems. For many reasons the assertion of such a purpose would be an act of some boldness anyway. The central aim here instead concerns the fundamental principles of such a theory, which essentially means that it has to deal with four particular complexes of questions: first, a precise terminology has to be elaborated to allow description of the essential aspects of formal, natural, and especially social systems. At this point it will already be crucial not to reduce the characteristics of

social systems to those of natural systems. Second, it is about the analysis of formal systems. This is supposed to demonstrate which methodological apparatus for the analysis of social systems is presently available, what knowledge about social systems can be gained from this analysis, and where open questions of research with regard to social systems appear. Third, it is about the exemplary demonstration that, and to what extent, characteristics of social (and to some extent also psychic) systems can indeed be mapped into formal models, and how they can be analyzed without making the mistake of reducing their content. Finally, it can be shown that we can already gain some new insights into social systems and their dynamics with this procedure, which will thereby open a new field of research.

In this context it is important to emphasize that we are dealing with the study of *dynamic* systems. When confronted with formal models most non-natural scientists usually think of logical structures, diagrams, etc., so what they have in mind are the static properties of what is being modeled. Here, however, we are dealing with the modeling of social dynamics. This is done for the simple reason that a static state is always just a borderline case of dynamics, as any physicist could tell. To work out the essential features of a really general theory of social systems one has to focus on the systems' dynamics. Here, appropriate computer programs - formal dynamic systems - become very important. In this study, thus, formal modeling always means: developing a dynamic model, implementing it into a computer program, analyzing the behavior of this formal system, and, finally, substantially interpreting the gained results.

The rest of this book is dedicated to working on these problems. It is, to stress it once again, not a new theory that claims to answer all sociological questions. It is rather the first step towards a theory that may be formulated one day by a sociological Newton. This cautious mode with which things are (usually) treated in the natural sciences is something that theoretical sociologists should definitely adopt.

Those who know about established mathematical sociology will probably realize that I will touch on this field only peripherally. This is certainly not because I think little of the achievements of that branch of research - on the contrary (see especially Fararo 1989 and 1997 on that). Theoretically, this study starts out from systems theory approaches and - in the actually sociological examples - the classical sociological theories; both are seldom done in the established mathematical sociology.[1] Technically, the study is oriented toward the new possibilities of

[1] One of the most important exceptions is Fararo 1989 who combines classical theoretical approaches with formal methods, similar to the attempt made in this study.

modeling that have emerged in connection with the exploration of complex systems, and develops them further to use them for the good of sociology. This is also seldom attempted, since mathematical sociologists instead use traditional mathematical techniques (see, e.g., Banks and Carley 1997). I do not think these techniques are sufficient for the comprehension of social complexity, which is why I have turned to the analysis of new possibilities. However, by their fruits ye shall know them, and therefore the reader has to judge if the innovations that I suggest can really bring something new.

1.1. SOCIOLOGY AND SYSTEMS THEORY

> "Zugleich kann man beobachten, daß das Forschungsfeld, das mit 'allgemeine Systemtheorie' bezeichnet wird, sich rasant entwickelt. [...] Die soziologische Theoriebildung könnte viel gewinnen, wenn sie an diese Entwicklung angeschlossen werden könnte." (Luhmann 1984, 15)[2]

As I mentioned above, the reason for starting my reflections primarily with the system concept - a term to be defined more precisely later on - is the following: rules of social interaction constitute a social system which have analyzed *sui generis* on the basis of their regularity. This understanding of the concept of a system is not necessarily the one that sociological systems theorists would assign to it; it is not my aim, however, to examine this particular problem more closely here. What is more important is the quoted suggestion by Luhmann, who formulates a program that at first glance seems to be quite similar to my outline above. It might help if we sort out to what extent Luhmann cannot achieve his own goals - at least not in the way he has chosen. For this reason, and to clarify my suggested approach - and, of course, also because of the outstanding effect Luhmann has had as a sociological theorist - I'd like to briefly sketch out Luhmann's program. It is exemplary for all attempts that pick up new terminologies while still sticking to established methods of the social sciences and the humanities, and thus also for all traditional attempts to develop social systems theories.

Since the publication of his main work "Soziale Systeme", Luhmann is known to have pursued his proposal of a "theory junction" with an overwhelming productivity. He was so successful that in sociology the term "systems theory" has almost become a synonym for Luhmann's theory. This cannot only be explained by Luhmann's productivity and his application of his "theory architecture" to almost every relevant social field. It has to be explained further by the fact that Luhmann seriously

---

[2] "It can be observed that the research field of general systems theory is developing rapidly ... Social Theory building could gain a lot if it could be connected to this field" (my translation, J.K.)

carried out his suggestion by explicitly using a major concept that was derived from a non-sociological context: the concept of autopoietic systems. By this means, he seemed to have carried out a theoretical junction that followed recent "paradigms" in systems theory, namely the natural scientific and interdisciplinary discourses of "self-organization", "synergetics", "dissipative structures", and "autopoiesis" - concepts from physics, chemistry, and biology. Luhmann seemed to have taken sociology to the level of the leading scientific disciplines - at least terminologically. The fact that a vast number of interdisciplinary and social science publications appeared on these subjects has a lot to do with the prestige Luhmann gained. On the other hand, it also increased his prestige in the process - a classic example of positive feed back, as it were.

However, there is room for doubt that a true junction to general systems theory has actually taken place by Luhmann's hand. What is overlooked only too often in the social sciences and the humanities - to the extent that they use systems theory concepts - is the fact that if natural science systems theories do adopt new "paradigms", which in itself is not a matter of course, they still follow the established traditions of theory and methods in the natural sciences. After all, natural science systems theories are not utterly new sciences, but are instead attempts to comprehend phenomena that were beyond the possibility of scientific analysis before, as a result of their "complexity". The sciences of chaos and complexity, which will be the subject of the next subchapter, have led to new terminologies and, above all, to improved methods of comprehending complexity. But still, they are mathematical sciences of nature, and that is precisely where they receive their relevance for interdisciplinary discourse. The characteristic formation of concepts, which is often adopted by interdisciplinarily interested members of the humanities and the social sciences, only makes sense in connection with a specific formation of theories and methods. To put the point mildly, it is at least questionable whether these concepts can be simply isolated from their contexts, and implemented into completely different contexts of theories and methods.

Looking at Luhmann's work from this point of view soon reveals a most elaborate and high-ranking adoption of several systems theory concepts into a completely different scientific culture, namely that of the humanities and the social sciences. Luhmann's theory clearly belongs to the classical and glorious tradition of the great social theories, and is in my opinion a worthy successor of such important preceding theories as that of Marx, Durkheim, Weber, Parsons, and others (including Habermas). As a consequence, it also belongs to a tradition of theory

formation that is primarily related to the humanities, whose strengths and weaknesses it shares. That is why Luhmann's theory has little in common with natural science systems theories. It is no accident that discussion of Luhmann's theory takes place almost exclusively among members of or social scientists oriented toward the humanities.

If Luhmann's suggestion of a theory junction for theoretical sociology is taken seriously, it is certainly not enough to use some basic systems theory concepts, which in any case have been chosen rather eclectically, and build up a new theory on that. One also has to show that at least basic parts of the theoretical and methodological context from which these concepts were taken can be implemented into one's own scientific approach. It should be obvious, especially to scientists with a hermeneutic understanding, that you cannot just adopt concepts without their contexts. So this study is about the question of how such a conceptual adoption can be made fruitful for sociology, without improperly reducing sociology to the natural sciences in accordance with the 'unified science' of logical positivism (cf. e.g. Apel 1979). In fact, I positively agree with Luhmann that with a theory junction sociological theory formation could gain a lot. First of all, however, one has to make sure what is actually being joined, i.e. what the genuine achievements and procedures of the other discipline are, and, above all, how it is possible to adopt at least parts of the contexts of theory formation and methods of the stimulating discipline.

It seems that Luhmann's concern has to be formulated more precisely. If you agree with his evaluation that theoretical sociology really could use some stimulation from outside (which I do), and if the possible candidates for such stimulation are mainly the recent (systems theory) developments in some natural sciences and branches of the computer sciences, then a sociology that sees itself as systems theory cannot just adopt a few concepts. Instead, it has to carry out two tasks for such a "junction": first, all essential concepts of extra-sociological systems theories have to be checked for their usability, and not just a few that seem to fit. Concepts, too, form a context, and one cannot arbitrarily pick some out of it. Second, it has to be analyzed to what extent the key concepts depend on the use of specific methods, and if they degenerate into simple metaphors without the adoption of these methods - which may be heuristically fruitful, but haven't really got any scientific content (Mayntz 1990). We shall see that this applies to practically every important concept. As a consequence, these methods must either be adopted, as far as possible and reasonable in sociology, or the concepts shouldn't be used at all. In the latter case, theoretical sociology would have to refer exclusively to its humanities tradition again.

Maybe this is exactly what some sociological theorists want. One reason for Luhmann's immense popularity is probably the feeling of being on the theoretical and conceptual level of the natural sciences without giving up one's own academic identity.[3] But unfortunately, you can never cross the river and remain on the old bank at the same time. If a sociological systems and action theory (see above) proceeds with methodological strictness, it must - unlike Luhmann - really leave the old bank, and will have to get involved in the uncertainties and adventures of the opposite side. This is not necessarily desirable, and in sociology there will always be research problems that can only be dealt with in the traditional way. What has become impossible, however, is to fall back on the reassuring remark that the "real" problems of theoretical sociology are beyond the possibility of formalization. It will never be as simple as that anymore.

1.2. THE NEW SCIENCES OF COMPLEXITY

> "What is being modeled by a computer is itself a kind of computer." (Churchland/Sejnowski 1992, 7)

The branch of theoretical sociology that understands itself as systems theory only partially takes note of the fact that in the eighties and nineties several research tendencies have developed which have a lot in common with the "paradigm of self-organization" and the traditions of general systems theory as established by von Bertalanffy; but that also go beyond this point in many respects, and have opened relatively new fields of research. There is no set name for these tendencies yet, and there probably won't be one in the foreseeable future: there are branches of research like chaos theory (Gleick 1987; Eve et al. 1997), which is already relatively established, and the theory of complexity, which partly developed from it (Lewin 1992; Waldrop 1992), as well as approaches like "computational science" (Churchland/Sejnowski 1992; Forrest 1990), or the experiments in connection with the "artificial-life" scene (Langton 1988; Langton et al. 1992; Langton 1994; Levy 1993). This list is certainly not complete. At first glance, these different research tendencies may seem very heterogeneous and somewhat confusing; but looking at them again, they have some fundamental features in common. These features are especially crucial for systems theory analysis. They can be characterized by the following concepts: (a) complex adaptive

[3] This holds true also for sociological systems theorists which orient themselves mainly toward the models of classical cybernetics.

systems, (b) bottom up approach, (c) computer as a model, and (d) interdisciplinarity.

(a) As regards content and subject matter, the relevant studies focus on the analysis of "complex and adaptive" systems. That entails the systematic inclusion of areas that were beyond the possibility of classical natural scientific research to date, and whose features reappear in many different phenomena (a traditional concern of general systems theory). A subject list made by Gell-Mann (1992, 8) to characterize the tasks of the famous Santa Fé Institute is best suited to show this:

> "Most of the work of SFI is on complex adaptive systems, systems that evolve or learn, such as;
> 1. prebiotic chemical evolution, including the chemical processes that gave rise to terrestrial life around four billion years ago; ...
> 2. biological evolution, leading through mutation and selection to the enormous variety of life forms on earth and also to the existence and evolution of ecosystems;
> 3. the behaviour of vertebrate immune systems, ...
> 4. individual learning and thinking in animals, including human beings;
> 5. human cultural evolution ...;
> 6. the global economy as a complex, evolving system ...;
> 7. the programming of computers to evolve, by mutation and selection, new strategies that no human has designed, for example for playing games."

The concept of "complex adaptive systems", which will be defined more precisely in the second chapter, refers to a fundamental characteristic: it concerns not only the analysis of processes inside (or of) a system, which is assigned the problematic attribute of complexity, but - in contrast to the above mentioned "paradigm of self-organization" - great importance is also placed on both the system's dynamics, i.e. its changes of state, and its evolution in relation to a specific environment. Such systems are characterized as "adaptive" because their self-organization processes - though highly complex - can only be understood as responses to their prevailing environmental conditions. This problem will be a subject for further examination.

(b) Methodologically, almost all of these new research tendencies follow a so-called bottom up approach. Langton's programmatic comment (Langton 1988, xxii) on the "Artificial Life" research program shows best what this is about:

> "Such systems contain no rules for the behavior of the population at the global level, and the often complex, high-level dynamics and structures observed are emergent properties which develop over time from out of all the local interactions among low-level primitives ... These emergent structures play a vital role in organizing the behavior of the lowest-level entities by establishing the context within which those entities invoke their local rules and, as a consequence, these structures may evolve in time."

A "top down" approach on the other hand, which is characteristic for the traditional techniques of simulation (cf. Klüver 1995, 102), starts out from the global behavior of a system, and usually analyzes only this. Of course, local interactions between the elements are presupposed here too, but instead of analyzing them in detail, they are usually summed up in highly aggregated equation systems. Such a procedure is especially applicable with "multi-particle systems" (which appear quite frequently in classical physics) in which the multitude of local interactions cannot be computed in detail anymore. In this case, however, it has to be taken for granted that it is actually possible to describe the global system behavior with global equations - a condition seldom met in the social sciences. Therefore, it is no accident that the option in favor of bottom up approaches is preferred by the new research tendencies characterized here: it results from the realization that really complex systems can only be recorded precisely if one starts out from their local interactions.[4]

Langton's quotation clearly marks the important point: to make no, or at least as few as possible, assumptions about the system as a whole, even if this is sometimes hard to maintain in research practice. In this approach, the system itself is not the starting point, but one of the results, *viz.* an emergent phenomenon. Therefore, the concept of emergence is as important for these research tendencies as the above mentioned concepts of complexity and adaptation. Furthermore, it follows from this claim that the option in favor of bottom up approaches cannot be seen as an ontological reductionism - however "natural scientific" all of these theories might be: it is explicitly taken into account that though the global system behavior is an emergent result from local interactions, it can influence them in reverse, and that the interactions of systems - global interactions, so to speak - also can and often have to be analyzed without recording the local interactions. Since the dynamics of global interactions often emerge in a "non-linear" way from local interactions, it is possible to say, in short, that the global interactions follow from the systems' "total" character. This is nothing mysterious of course, it is just a consequence of the complexity of systems.

(c) Theoretically, as well as methodologically and technically, computer modelings are at the heart of all these research tendencies. This

---

[4] The analysis of probably the first ever systematically examined system, namely the planetary system by Kepler and Newton, already shows this difference very clearly: Kepler's laws - the determination of the planetary orbits as ellipses - are a typical top down approach, because they describe the system behavior only on a global level. Newton's laws of gravitation on the other hand, lay down the interactions between the system elements - planets and sun - and produce Kepler's laws as a logical consequence. Only the bottom up approach of Newton's equations, according to the common view in physics, can fully explain the planetary system.

does not only mean that the construction of computer simulations naturally belongs to their methodological tools. The use of such tools has become typical in everyday research work in all natural sciences, and has begun to catch on in the social sciences, too (see below). It means instead that in many respects the computer itself (or else particular computer programs) is taken as the model for what is being researched, and that the respective subject is characterized as a computer in a more than just metaphorical way of speaking. This is obvious with Artificial Life research; it is not so obvious when such a complex system as the brain is described as a computer - the brain, not the mind. Correspondingly, one can assume that the subjects of the above list are explicitly or implicitly seen as computers by Gell-Mann, which is not a trivial assumption. To make this assumption a little more plausible, and to soften its well-known imagery, I'd like to sketch out some of Churchland and Sejnowski's argumentation as an example:

Churchland and Sejnowski are perfectly aware of the fact that the computer metaphor is first of all quite hackneyed already, and has furthermore led mostly to a confusion of concepts and to a thirst for pseudo-exactness. This is particularly true when the technical construction of certain computers is taken as a yardstick. By contrast, they begin by giving a very general definition of what a computer is, without regard for its specific physical realization:

> "We can consider a physical system as a computational system when its physical states can be seen as representing states of some other systems, where transitions between its states can be explained as operations on the representations. ... That is, the physical system is a computational system just in case there is an appropriate (revealing) mapping between the systems' physical states and the elements of the function computed." (Loc. cit., 62)

So, what is meant is the following: a computer is a physical system whose states, and whose transitions between these states, can be mapped into other systems, which it represents. These other systems are primarily the formal systems of mathematics: the transition of one state of the computer (the physical system) to another is, on the level of the formal system, the computation of a new value of the function from a given one. In this sense, the brain "is" a computer, because the transitions of one state to another - from one configuration of the neurons and synaptic links to the next - can be mapped into the operations, or "computations", of an immensely complex formal system (for appropriate neurophysiological examples see Churchland/Sejnowski, 1992). Accordingly, the famous stones at Stonehenge "are" a computer, because:

> "The system of Aubrey holes at Stonehenge computes eclipses of the sun by dint of the fact that its physical organization and state transitions are set up so that the sun stone, moon stone and nodal stone land in the same

> hole exactly when an eclipse of the sun occurs." (Churchland/Sejnowski, 66)

Exactly in this sense, Dewdney (1988) demonstrates how bundles of spaghetti or knottings of ropes can be seen as (analog) computers.

So far we have only talked about physical systems. This seems to rule out the possibility of including "immaterial" systems like mind or society into this paradigm - that is to say, to think of them as computers in the above mentioned way. There is no reason to assume this, however, since the concept of the "virtual" computer, in contrast to physically real computers, has been a set and well-defined technical term in the computer sciences for a long time (see below, chapter 3). In a generalization corresponding to the above definition, any system, whether physically real or immaterial, can be understood as a computer - physically real or virtual - as long as there is an appropriate mapping of it into formal mathematical logical systems.

Such a characterization of the system can, of course, still be taken as a metaphor, and one can argue whether much is gained with it *per se* (cf. e.g. Harnad 1994). The reason why it is more than just a metaphor for the above mentioned sciences of complexity is that it results in a specific and quite effective methodological consequence: If "real" systems can be interpreted as computers, then they can also be mapped into formal systems that are usually described as computer programs, for example cellular automata or neural nets. This again leads to the methodological conclusion that it is useful and necessary to analyze appropriate computer programs for their own sake, and to put equal emphasis on their analysis in the study of complex systems. That is why Gell-Mann's list contains the construction and analysis of complex computer programs as a specific and crucial point. We shall see that this computer paradigm, which should rather be called a "disciplinary matrix", because of its very substantial methodological consequences, also leads to a new view of the concept of computer simulation.[5]

(d) As regards the sociology and history of science, those new sciences are on the one hand part of the culture of the natural sciences; on the other hand they are to a high degree interdisciplinarily organized. A list of publications of the Santa Fé Institute, where studies on physical-chemical and biological systems appear next to publications on linguistic

---

[5] Contemporary neurophysiologists often emphasize that the brain is not a computer, since it works on fundamentally different principles than the customary computer programs, or its hardware, respectively (e.g. Edelman 1992). According to the above definition, however, it is quite certain that the brain is a computer, because no neurophysiologist (nor even Edelman) would doubt that a mapping between the brain and a highly complex formal system is possible.

systems (Hawkins/Gell-Mann 1992) and economical systems (Anderson et al. 1989), shows this quite clearly. The concept of interdisciplinarity needs to be stated more precisely here: the fundamental approaches and research methods originate in an extended natural science culture in the sense described above, *viz.* as an analysis of its subjects *via* the construction of formal models and computer implementations. Special subject contributions of particular disciplines are integrated as far as they are able to contribute to the exploration of "complex adaptive systems". Of course, these new research tendencies share their fundamentally interdisciplinary orientation with traditional general systems theory. What is new, however, is that they put the general paradigm in concrete form in the ways described here.

Since the first analyses of chaotic systems in the sixties, the use of computers has played an important role. It is well known and has been described many times (e.g. Gleick 1987) how the first results of the behavior of chaotic systems and the phenomena of so-called attractors (see below, chapter 2) came about almost accidentally: the meteorologist Lorenz had left out a few decimal places in the computer calculation of a climatic model, which led to crucial changes in the behavior of that model as the simulation went on for some time - one of the characteristic phenomena in connection with chaotic systems. This great sensitivity of chaotic systems to slight changes of particular parameters and other features could only be recognized and studied by creating computer modelings that were more and more precise. In other words: a very large part, if not the largest, of the entirety of chaos research, whose meaningfulness is undisputed among physicists, took place "inside the computer" - to such an extent that critics occasionally doubted the reality of some of the results and described them as computer artifacts. This holds true even more for the theory of fractals, whose popularity is undoubtedly due to the fact that although the mathematical theory behind it is rather elaborate, it can be demonstrated quite easily with a computer (Mandelbrot 1982; Dewdney 1990).

Of course, computer simulations have not simply replaced the classical natural science methods, namely experiment and mathematical theory formation. But since the research into complex systems in particular more and more often reaches areas in which classical experiments cannot be carried out (e.g. weather, ecological systems, biological evolution), and classical mathematical procedures do not apply, because there are no simple analytic solutions for these problems, the construction of computer models is often the only way to obtain precise theory formations. Hence, all these new sciences are computer sciences in several respects: they take the computer (in the sense

described above) as a model, and for the most part they also take place inside the computer.

The role of the techniques of computer simulation, which are well established in the natural sciences and parts of economics, usually lies in modeling simulated systems on the basis of equations. This formal model - the equation system - is then implemented as a special computer program and run with empirically ascertained initial states (initial values) and boundary conditions. One application of the equations to the initial states produces new values, on which the equations are applied again, and so on, until a certain outcome is taken as a "final" result (for practical reasons). Thus, simulations are just the constant "iteration" of equations by the simulation program, i.e. the continual application of the equations to themselves. If the equations are described as the function(s) for the simulated system, then a simulation is a constant "iteration" of the system function(s).[6]

Basically, this iteration of system functions counts for all computer simulations of dynamic systems - there are, of course, restrictions, which will be dealt with in chapters three and four. Admittedly, these established techniques of simulation have a level of abstraction that often cannot be executed in the social sciences - and not only there: epistemologically, the equations characterizing a simulated system naturally do not have much in common with the original system. They arise from observing the behavior of the system or its elements, and are confirmed by comparing the simulation with the real system dynamics. Basically, this is what the mathematical natural sciences have always used as their method of experience, and so it is not completely new. That is why even today the method of simulation is often rated as secondary by established natural science theorists: it is a necessary procedure if elegant and analytically simple solutions are not available, as it is the case with many classical problems, but there is no "proper" theory until there is a classical solution for an equation system.

The examination of highly complex systems, especially in the fields of artificial intelligence (AI) and artificial life (AL), has somewhat changed this procedure. Since it is absolutely impossible to describe such systems in a traditional mathematical way, the starting point here is to analyze the *fundamental* characteristics of the system that is to be simulated, e.g.: biological evolution by mutation and selection in the program "Tierra" by Ray (Ray 1992); the reproduction of organisms by cellular automata - John von Neumann's famous problem (Langton 1984); or intelligent

---

[6] Actually, this is only valid in the case of equations whose solutions cannot be analytically determined so that they have to be approximated. We can abstract from this point here, however.

systems' problem solving behavior as in classical AI. After that, a formal system, *viz.* a computer program, is constructed to exhibit these characteristics *in its own behavior* as similarly as possible. The rules of the computer program do not necessarily have to mirror those of the simulated system, but - in accordance with the famous Turing Test - the formal system shall first of all show a similar behavior; in a way, it shall be the formal equivalent of "life", "intelligence", "reproduction", "evolution", etc. That some people then talk about these programs as being actually alive or intelligent is another matter, which is of no interest here.

In this procedure, simulations are obviously not just secondary, but are the essence of the entire endeavor. Consequently, their analysis does not primarily convey something about single concrete systems, but about the relevant *system type*. It is primarily this procedure that I will use in the following chapters to study the logical characteristics and the dynamics and evolution of social systems in general. To avoid any misunderstandings: the two ways of using simulations do not contradict but rather complement each other. First, which of these two options one chooses is a question of research interest - depending on whether one is more interested in single concrete systems or in the analysis of system types. Second, there is the question of the object of research - does it actually allow classical mathematical descriptions? If not, the second way has to be chosen. In the next subchapter I will put together some arguments to support the claim that the classical objects of the humanities and the social sciences do presently not, and probably never will, allow the traditional means of mathematical description. A social physics, as dreamed of by Comte, will probably never be possible in the classical style; but there are now initial signs of what the new ways would have to look like (see also Casti 1997).

I have sketched out the new sciences of complexity here only formally and from a rather methodological point of view. In the next chapter I will give more attention to their contents. There's one point, however, that may not be withheld at the end of this sketch, namely the impression one gets all too distinctly from reading numerous research reports, summarizing monographs and anthologies, and articles in scientific journals: there is a prevalent and partly excited pioneering spirit, the feeling of opening new pages in the book of science, of finally being able to tackle centuries-old basic problems with precise means, or at least to bring them closer to a solution. The texts in question sometimes remind me of the great Soviet psychologist and linguist Lurija's beautiful concept of the "romantic science" (Lurija 1993).

That mood is expressed quite aptly by Churchland and Sejnowski, and I would like to end this sketch of a new departure in the systemically oriented sciences with their words:

> "Nevertheless, the hunch that exciting things are in store is difficult to subdue, and the thrill of discovering what we are and how we work is luring more and more students into the fields. ... Withal, it is a remarkable time in the history of science" (loc. cit., 16).

Maybe I can communicate some of the fascination I have felt myself quite often while working in our research group at Essen.

1.3. SOCIAL AND COGNITIVE SYSTEMS: THE SYSTEMIC OBJECTS OF THE HUMANITIES AND THE SOCIAL SCIENCES

> "Die Platonisten und andere Vernunftspuristen neigen dazu, über ihrem Instrumentarium der theoretischen Vernunft ... die Instrumentarien der dynamischen Wissenschaften (Zustandsräume, Trajektorien, Attraktoren, um nur ein paar Schlagworte zu nennen) zu vernachlässigen." (Schnelle 1994, 119) "

(Platonists and other partisans of pure reason tend ... to neglect the instruments of the dynamical sciences (state spaces, trajectories, attractors, to name only a few slogans).

In the context from which this quotation is taken the linguist Schnelle argues against an overly rigid emphasis on the logical-symbolic procedures that have been dominant in linguistics since Chomsky. This is certainly not a problem for sociology, because an introduction of formal methods similar to linguistics has not often taken place there. It is not irrelevant, however, that on the part of linguistics this emphasis is increasingly seen as a chance to take advantage of the new sciences of complexity.

Since it is my goal - in accordance with Luhmann's program - to connect theoretical sociology to these new systemic sciences, as opposed to merely adopting a few of these fields' concepts, it is especially computer-centered procedure that must be adopted (an orientation that has been taken as a matter of course in linguistics, especially computer linguistics, for a long time.) First of all, however, one has to make sure of the objects - the systems - that are relevant for the "Kultur- und Kognitionswissenschaften" (cultural and cognitive sciences) (Schnelle). If my remarks in the last subchapter are correct in asserting that simulations in the humanities and social sciences have to be primarily of the second type, i.e. that the crucial point lies in the construction of programs that show in themselves the fundamental characteristics of the

simulated systems, then the first question that arises concerns the nature of these characteristics.

In a simplifying way, which is inevitable here, we might say that the humanities and the social sciences have principally taken three problem areas from classical philosophy as their genuine subjects: language, mind, and society. The fourth classical problem, namely the question of nature, has long since become a monopoly of the natural sciences. My concern is that we are dealing with particular system types, which are to be described as linguistic, cognitive or psychic, and social systems. In my opinion, the concept of the "cultural and cognitive sciences" expresses this quite aptly, though without any connotations regarding systems theory. So what are those fundamental and, especially, common characteristics of these systems?

First of all, they are obviously self-organizing. For the most part, one can follow Luhmann's argumentation regarding social systems (1984) and cognitive or psychic systems (1985), though for various reasons one should not follow his specific terminology. In this case, it makes no difference if you speak of autopoietic systems, as Luhmann does, or instead use the more neutral and more common term of self-organization, as I do. The significant point in both cases is that systems are subject to a special kind of internal dynamics, which makes them "operationally closed". In other words: the system's operations, and especially its formation of particular structures and "organizations", do not occur regardless of its environment; but the environment cannot influence the system-specific way in which these operations happen, unless the system collapses and/or merges with the environment, and thus loses its identity. These general remarks may be enough for the moment, since the relevant concepts will be defined more precisely in the next chapter.

Since Luhmann has focused on this aspect of social and psychic systems in numerous analyses, I can refer to the essential features of his argumentation for these two system types. The corresponding analysis of language and linguistic systems has been less detailed, because the theories of dynamic systems (see above) met with little response in linguistics. One can assume, however, that the new linguistic theories since Chomsky, who put a strong emphasis on the autonomous character of language, can be appended here quite well: language can be understood as a self-organizing system insofar as the grammatical, morphological, and phonetic rules of speech production follow their own logic that can - and sometimes must - be analyzed as an object *sui generis*, and that cannot be *directly* influenced or controlled by its social or psychic environment. In this general sense it is possible to see language as a self-organizing system.

Second, these systems are adaptive. I describe systems as adaptive if they are able to orient themselves to an environment without losing their identity. Since they are also self-organizing, there cannot be a direct influence of the environment on the system, which would determine the system's own logic - in which case it wouldn't be its 'own logic' anymore. I mean instead that these systems can vary their internal rules of operation in such a way that they can change the principles of self-organization within certain frames while keeping their fundamental self-organizing character.

The theorists of self-organization have given the problem of adaptivity relatively little attention; and the same could be said of the theorists of adaptation concerning the problem of self-organization. In any case, it should be evident that in this sense, which is still rather general, linguistic as well as psychic and social systems are adaptive: linguistic systems vary their grammatical, phonetic, and morphological rules - to say nothing of their orthographic rules - depending mainly on their social environment; psychic systems change the logic of their thought production depending on their developmental stage and the demands of the situation; and social systems change too, depending on the prevailing environmental conditions. In the fourth chapter several examples of this latter case in particular will be given.

There is another characteristic that actually goes with the adaptivity of the systems that are analyzed here: it is the fact that they model their environment, i.e. that they carry out their adaptation by rearranging their self-organization on the basis of a certain picture of their environment. In the fields of language and mind this claim has often been debated; within sociology, the problem has been discussed especially in the sociology of knowledge.

Accordingly, these systems are evolutionary in the sense that they first build up the principles of their self-organization themselves, and then - usually in co-variation with their environment - execute them within the bounds of their possibilities. Piaget's well-known model of development can be used here as a classic example for the human ontogenesis (e.g. Piaget 1975): the mind develops in the course of assimilation and accommodation so that its fundamental schemata are gradually built up, which then - by means of self-organization - control the actual cognitive processes. Something similar certainly holds true for the evolution of languages and - as we shall see later - for social systems.

Furthermore, the systems are able to learn in a formal sense of the term (cf. especially Bateson 1972). According to Bateson, systems are able to learn if they can at least (a) respond differently to repeatedly occurring situations if their behavior wasn't successful the first time that

situation occurred, and (b) vary the learning strategies themselves in a certain way by "second level learning." I think that these criteria apply perfectly to psychic and linguistic systems - see my remarks above on adaptivity and evolution. There will be a separate analysis for social systems and computer programs in chapters 2, 3, and 4. All systemic characteristics that are listed here also apply to biological systems like non-human organisms, animal populations, etc. The peculiarity of the systems analyzed here is basically due to the radicalness and extent of the degree of freedom with which these characteristics are exhibited. This is especially true for a further feature, which elucidates why the cultural and cognitive sciences have established themselves as an independent scientific field alongside the natural sciences, and why all efforts at the standardization of science had to fail: it is the capability of modeling itself within itself, i.e. to be able to "get an idea of itself", particularly by means of the system operations themselves. I'd like to call this process self-modeling or self-simulation. Since this aspect is rather neglected in systems theory literature, I'd like to look into it in more detail.[7]

Presumably, it is quite clear what all this means for linguistic and psychic systems. It is known that for the part of language which is usually and misleadingly called "natural language" by logicians who want to set it off against formal languages, it is no problem to model a language by linguistic formulations - linguists and language philosophers basically do this for all their professional life. Therefore, the "natural language", i.e. the scientifically or in some other way enriched colloquial language, is often described as the supreme meta language (e.g. Apel 1979), which stands above all other languages because it can speak about these and also about itself too. This is where the practitioners of logical semantics saw the source of the semantic paradoxes (Tarski 1956). A psychic system, i.e. the mind, can also be defined by its capability of thinking about itself, in the way that sequences of thoughts create a state of consciousness which is a self-conception of the mind.

For social systems this characteristic is more difficult to describe. Sociologists who are somewhat familiar with Luhmann might think of a self-description of society, e.g. in the form of specific semantics. However, this only makes sense if you agree to Luhmann's terminology, which is derived from the humanities and is rather unsuitable for a

---

[7] Sometimes this is treated under the name of "self-referentiality" or - in Luhmann - "Reflexion". The paradoxes of semantic self-referentiality have become famous (e.g. Hofstadter 1985 and 1986). Yet this was usually not about the special ability of self-modeling but about more general forms of self-referentiality. In neurophysiology however, the problem seems to be playing a more and more important role under the subject "brain and mind" (cf. Roth 1996): presumably, this is exactly what consciousness is, namely the self-modeling ability of the system "brain".

systems theory that strives for precision in a formal sense. If self-modeling is taken as a substate of the system, which like any other state is set into effect by self-organizing processes, and therefore is the result of system operations, and which constitutes a model of itself for the system, then society's capacity for self-description has to mean something other than what Luhmann's terminology suggests.

There is a very instructive example, which results from Geertz' (1972) field researches on the island of Bali. The issue is basically the following: Its subject is the ritual of organized cockfights on the island of Bali. Geertz meticulously describes how the "social matrix" of the Balinese society is simulated or reproduced in the social subsystem 'cockfight', particularly by the placing of bets: people on a lower social level can only take part in small bets; the higher ones participate in high bets; only very respected persons are involved in the center of events; no bets are made against members of one's own village community, etc. Most importantly, only men take part in these rituals. This fact, which is only mentioned *en passant* by Geertz, reinforces, of course, the character of the entire event as a model of a patriarchally organized society. Geertz reconstructs the cockfight as a self-conception of Balinese society and explains it like this: "... it is a story they tell themselves about themselves" (1972, 26). This and other examples by Geertz concern the canalization of aggressive behavior by means of modeling simulations; corresponding parallels to professional sports are obvious.

The social system of modern science has often been described with the concepts of "trial-acting" and "play": possible actions concerning nature and society are tested experimentally, and can be undone if necessary, before they are carried out in social and industrial reality. In fact, in the sense described above, one can take the social system of science as a modeling of the society by itself with a simulation of the social matrix - the hierarchies of relevance of the various parts of society reproduce themselves in the hierarchies of relevance of the scientific disciplines; social inequality reproduces itself in the social gradation of the social system of science, and so on. There is no need for further explication of the fact that science is taken at least as serious in modern societies as the rituals in Balinese society described by Geertz. What these examples tell us is that a social system's modeling of itself is obviously a universal characteristic of social systems. As a process, self-modeling is formed by the system operations - social interactions according to specific rules; as a result, it is a substate of the social system.

Evidently, systems of this type - self-organizing, adaptive, evolutionary, able to learn, and self-modeling - are the most complex systems actually known.

The reasons for this fact will be analyzed in more detail in the following chapters; furthermore, this list of features is probably not even complete and might well be supplemented by further research. The systems' high complexity makes it understandable from the point of view of systems theory why they have persistently eluded all attempts at formalization - to say nothing of the difficulty or impossibility of systematic experiments. Schnelle (loc. cit.) rightly remarks that linguistics, too, is still far from having the status of a precisely explanatory science. The issue is rather whether in the cultural and cognitive sciences, whose objects tend to be overcomplex, there can be theories after the model of natural sciences at all. In other words, and referring to the central subject of this book: is it actually possible to represent such systems in formal models, which could serve as a basis for accordingly complex computer programs - that is to say, programs understood as formal systems that belong themselves to the type described here?

One of the essential claims of this study is that this question can be answered affirmatively. There really are such programs, and we already work with quite successful prototypes of this level at the University of Essen. Naturally, these are programs that (have to) differ to some extent from the common understanding of computer programs as far as it rests on the idea of computer programs as fixed rule systems - algorithms - which do exactly what they were instructed to do. Such a conception was absolutely appropriate until the eighties, and consequently even a distinguished mathematician and physicist like Penrose (Penrose 1989) thought on the basis of this understanding that the possibility of something like artificial intelligence was completely out of the question. Here, however, we will talk about programs that show characteristics - though still fairly elementary - that are essential for linguistic, psychic, and social systems.

The fact that systems of this type can principally be mapped into formal models and consequently into computer programs - therefore the systems can themselves be interpreted as computers according to the definition of Churchland and Sejnowski - immediately raises the question of what kind of knowledge can be gained about these systems by such programs. As regards content, this question cannot be answered before the end of this book - and even there the answer will naturally be a rather general one. As regards the theory of science, we can already anticipate some results from the systems' characterization - in contrast to natural science theories and laws:

As I mentioned earlier, the new sciences of complexity are legitimate members of the natural sciences, regardless of their interdisciplinary inclinations, and regardless as well of the fact that economists and other social scientists are increasingly interested in these areas of research.

This natural scientific orientation finds expression in most of the participants' scientific biographies: most of them are mathematicians, computer scientists, physicists, chemists, and biologists; physicists especially are distinctly predominant. And so it is no surprise that the search for regularities of complex systems has almost always consisted - and still consists - in tracking down quasi-natural scientific laws, which apply equally well to turbulent liquids, to ecosystems, and to the turbulence of stock markets (an example that is repeatedly given). A short example will illustrate why this sort of perspective offers a much too simplified view on problems of the cultural and cognitive sciences:

For didactic reasons, namely as an exercise program for student work groups in a seminar on systems theory, Jörn Schmidt and I constructed a cellular automaton (see chapter 3) to simulate a simple predator-prey system. Our intention was that the students would compare the bottom up approach of a cellular automaton's modeling with the top down approach for the same system on the basis of the equations of Lotka and Volterra, for which we also used a simulation program (Bossel 1992). After having experimented with the cellular automaton ourselves for some time, it coincidentally and to our surprise turned out that on specific parameter values (reproduction rates of predator and prey) the dynamics of the system presented itself in the form of spiral waves: both predator and prey populations developed spiral shapes, which winded around each other to the left and to the right. Once the spiral patterns turned up, they remained stable in the sense that the system developed spirals that disappeared and made room for new ones in all simulation steps we carried out - in a word: there was a dynamic stability in the form of spiral dynamic patterns.[8]

What astonished us about this unintended result is the fact that such spiral patterns can occur in quite different kinds of dynamic systems, e.g. in the formation of galaxies, in chemical solutions, and in catalytic processes, to name just a few. Gerhard/Schuster (1995) give several examples, some of which also resulted from simulations with cellular automata; the circular and spiral waves of the famous Belousov-Zhabotinsky reaction have been a classic for a long time. One can easily realize why spiral patterns are one of the simplest forms of stable dynamics in dynamic systems that consist of two interacting element types, which is why they appear quite frequently; this occurs in processes

[8] After we had completed the program and the students had worked with it, I discovered that the American computer scientist and editor of the computer column in the "Scientific American", Dewdney, had already constructed a cellular automaton for the simulation of predator-prey systems a few years earlier. It is certainly no accident that the rules drawn up by Dewdney were almost exactly the same as ours.

that are characterized by diffusion-like interactions. Therefore it seems reasonable that theorists of complexity see in these and similar pattern formations universal regularities of complex system dynamics, and search for such laws - the simple rules behind the complexity of appearances - in other complex systems too. Our incidental results could be interpreted as indicating that in this respect ecosystems are similar to galaxies and chemical reactions.

However, we didn't believe for a second that we had made a new important contribution to the discovery of universal pattern formations. For one thing, the spiral patterns in our program only ensue from very specific parameter values or proportions, and for the most part they do not appear. Being non-biologists, we do not know which genera of predator or prey correspond to our reproduction rates anyway. So it might well be that the formation of spiral distribution patterns of the population always occurs within specific predator-prey systems, but this is certainly not necessarily the case. Establishing that this *is* the case would require in-depth biological research. Furthermore, our program is too simple to do justice to the ecological reality of a predator-prey system anyway; this is also true, by the way, for the equations of Lotka and Volterra I mentioned earlier. The program allows the examination of some of the basic behavior of specific dynamic system types, according to the research methods described in chapter 1.2; precisely by virtue of the fact that it is so simple, it is very transparent, and I will come back to it to illustrate various problems. The behavior of real ecosystems is a totally different question.

Finally, however, being natural and social scientists we were very aware of the fact that those complex systems we were dealing with above followed even more complicated dynamics. In our program, the initial rules and parameter values do not change, and so the described dynamics ensue by formal self-organization. This is an appropriate simulation approach for real physical systems, which do not change their regularities, like systems in physics and chemistry; and it is highly likely that with more complex programs one can find general regularities, e.g. in the form of pattern formations. But with regard to learning and adaptive systems, which can vary their rules and parameter values, the situation is totally different. Here it is not possible to analyze from which simple regularities the system dynamics have ensued. Instead one has to ask on a deeper level what changes the system has carried out in its rule structure to create the specific dynamics that can be observed. In a certain sense, the question about the regularities behind the regularities arises, or to put it in logical terms: the meta-rules above the rules and their combinations with each other. Obviously, an additional level of abstraction has to be introduced to at least partly meet the requirements

of the systems described above. Bateson in particular mentioned this point several times when he emphasized the necessity of thinking in logical types, even if he could not yet know about the possibility of computer programs with the needed capability.

This problem was often given too little attention by the theory of complexity, which is dominated mainly by physicists and mathematicians. Instead, the usual assumption was and still is that one has to search for regularities in the way described here only on one logical level. The Santa Fé Institute, for instance, has organized research programs to investigate the rising and falling of prehistoric societies in accordance with the natural scientific approach - an approach that in this case reminds one of the obsolete biologistic culture theory of the almost forgotten Oswald Spengler (Gumerman/Gell-Mann 1994). That might also be the underlying cause of the present talk of the "crisis of complexity" (Horgan 1995): the problems not only turned out to be more difficult than was expected during the first euphoria, but also and above all the attempt to apply natural scientific thinking almost directly to cultural and cognitive science areas - though with some methodological subtlety - did not consider the logical difference between the particular types of complex systems. It seems that the cultural and cognitive sciences' independence in comparison with the natural sciences can be justified and proved reasonable particularly from a standardizing approach like the general theory of dynamic systems.

Regardless of these necessary restrictions, it has to be stressed once again that sociological systems theory - and theoretical sociology in general - could be stimulated to a high degree by the theoretical models and technical procedures of the new theories of complex systems' computer simulations. It is just the difference from the likewise complex natural science objects that clearly shows the peculiarity of the (social) systems that we are interested in here, which can then be tackled as a specific research problem. By taking up the conceptual and methodological arsenal of the natural scientific theory of complexity practically and not only as an eclectic conceptual borrowing, a theory of social systems will precisely *not* become an appendage of the physics of dynamic systems, but is entitled to take the peculiarity of its objects as a legitimate justification for its independence. As a result, I think that the natural scientific theory of complexity in particular, with its occasionally imperialistic claim to dominance, could learn from other disciplines too, namely from the humanities and social sciences, which are often looked at rather contemptuously. This is where the problems of systemic complexity actually lie, and the natural sciences can offer suggestions, but certainly no solutions.

. Additionally, recent attempts to use the simulation methods of AI and AL in sociology (Gilbert/Doran 1994; Gilbert/Conte 1995) show how important a theoretical orientation to the natural scientific theory of complexity can be. Almost all of the programs presented there follow the theoretical orientation of methodological individualism: the programs, some of which are impressively constructed, model how emergent order patterns of groups, prehistoric tribal societies, and single organizations develop from the specific actions of individuals - often in connection with the rational-choice approach. What is left out of consideration, however, in most cases, is the fact that social patterns also retroact on the single actor, and that they develop a large degree of originality in contrast to him. That is something that has always been seen better and more deeply by the classical social theorists, including Luhmann. As a consequence of this theoretical deficiency of the merely individualistic approaches, Gilbert pleads for an orientation to the categories of the modern science of complexity.[9] It is my conviction that this can impart its more precise conceptual and methodological qualifications to the great tradition of classical social theory, which it strongly needs for its genuine task: to analyze society in its peculiarity as well as its similarities to other complex systems.

[9] One sometimes gains the impression from some of these projects' studies that they had reinvented the sociological wheel. When for example Conte and Castelfranchi (1994) suddenly discover under the title "Mind is not enough" that beside the consciousness of social actors there are also "objective", namely consciousness-independent relations, they are obviously not aware of the fact that large parts of sociology have dealt with exactly that problem for a long time.

2. STATE, EVOLUTION, AND COMPLEXITY: BUILDING BLOCKS OF THE THEORIES OF COMPLEX SYSTEMS

In the first chapter I repeatedly referred to the terminology of the sciences of complexity, which is the collective name I would like to apply to the variety of theoretical approaches in this area. I have not yet precisely defined the specific concepts with which I will work, which is what I would like to do in the present chapter, particularly with regard to the concepts that characterize the systems that interest us here. We shall see that in some important aspects we will have to go beyond the terminology of the natural sciences. The reason for this lies in the already noted peculiarity of the systems that cultural and cognitive sciences are dealing with. Concepts that are already established in the natural sciences, such as trajectories, state spaces, non-linearity, mathematical chaos, etc., will be used without further definition. There are numerous introductions and handbooks on such topics already. The concept of the attractor, however, will be described in more detail, because of its importance for our topic.

2.1. GENERAL CONCEPTS

We can define a system in the customary way as a set of interacting elements. The rules or regularities that determine these interactions characterize the system's dynamics, which is just a sequence of the system's states, generated by rules. The rules that determine the interactions, and with that the dynamics of the system, can also be understood mathematically as a function that maps one state of the system into the next. If that function is linear in a mathematical sense, we can speak of a linear system; otherwise, it is a non-linear system. Since the rules determine the transition from one state to the next, they can also be described as the transition function of the system. It shall not be denied, however, that the mathematical proof of linearity or non-linearity is not always easy to furnish in the case of rule systems that cannot be plotted directly as linear or non-linear functions (for the case of cellular automata see e.g. Kari 1992). Graphically, the dynamics of the system can be illustrated with trajectories. These are curves in the system's state space - that is to say, the set of all the points that define a possible state of the system. These trajectories describe real states of the system as a subset of all of its possible states. The state space is defined by the logically independent variables that are essential for the characterization

J. Klüver, *The Dynamics and Evolution of Social Systems*
© Springer Science+Business Media Dordrecht 2000

of the system - such as, for example, the number of predators and prey in a predator-prey system.

According to the definitions above, we can now simply characterize system-environment relations as rules that determine the interactions of (some or possibly all) system elements with specific parts of the environment. Non-linear systems are of special interest for the theories of dynamic systems - and not merely because of their relation to chaos theory and the theories of complexity. What is especially important is the question of whether the trajectory of the system shows *attractors*. This means that if one observes the trajectories of non-linear systems in the state space, the following effect can occur: for a certain time, the trajectory seems to proceed completely irregularly; after a while, however, it turns out that the trajectory does not leave a particular area of the state space (or does so only rarely) and "narrows" to one point in it. One could say that this point "attracts" the trajectory. Such a point of attraction, which holds the trajectory close to itself for the remaining course of the system, is called a (point) attractor of that system (or of its trajectory, to be more precise).

This point can be defined in more detail: the trajectory results from the fact that the rules of the system are continually applied to each achieved state. If we call the totality of rules the "system function" or "transition function" of the system, then the trajectory results from the continual iteration of that system function, that is, from the constant application of the function on itself. So if Z_1 is the initial state of the system (the "first" point in the state space), f is the system function, and A is a state that results as

$$A = f^n (Z_1)$$

where f^n is the application of f to itself n times, then A is an attractor of the period 0, or point attractor, if for any $m > n$ it is true that $f^m (Z_1) = A$.

This means that the trajectory stays at one point in the state space, which it has reached by the above equation, and does not leave. If a state is reached again after the realization of k *additional* states, and the trajectory permanently repeats this sequence of states, we get an attractor of the period k. Thus, an attractor of the period k does not define a single point in the state space, but a basin of k states, lying on a closed curve in the state space. Attractors of finite periods are usually called "simple" attractors.[10]

---

[10] As with other concepts too, the terminology is not entirely unambiguous here: Sometimes a point attractor is said to have the period 1. Accordingly, states that are reached again after k other states are attractors of the period k+1. Since the important point in the above definition is the number of *different* states between the reaching of an attractor, I will use the above definition here.

If the application of f on itself happens to produce a linear function F, e.g. ff(Z) = f(Z) + f(Y), then F obviously constructs a straight line in the state space. In this case, there is of course no attractor, because the (linear) trajectory constantly proceeds in the same direction. However, this can only happen if f itself is linear, which means that we are dealing with a linear system.[11] Thus, only non-linear systems can have attractors (except in case of an identical system function).

The concept of the attractor itself is relatively new, and has actually become relevant only since research in physics began working in connection with the theory of chaos. Mathematically, however, the concept has been established for a long time. It was introduced by Hilbert sixty years ago under the name "*Eigenwert*" (eigen value). However, I will use the term "attractor" in the following because this term is used exclusively in the context of the theory of chaos and the new theories of complex systems, and also because it is the more general one.

Systems can also have different attractors. In this case, only the initial state decides on which attractor is adopted by the system. This point will be set out in more detail with the formal systems of cellular automata and Boolean networks in section 3.1. Thus, attractors are a property of the *rule system*: only the rule system decides how many and which attractors a system fundamentally has. A "strange attractor" arises when the trajectory of a system runs through all points of a particular space segment in the state space. These points (infinitely many, in principle) are parts of the strange attractor. In this case, the behavior of the system's trajectory allows at best quite short-range prognoses.[12]

Now it is possible to define the concept of the chaotic system precisely: A (non-linear) system is called chaotic if it has either no attractors at all, or only strange attractors. Thus, a chaotic system has

---

[11] This results from the defining capacity of linear functions: For each linear function it is true that ´f(a +b) = f(a) + f(b). Of course, this goes for any number of summands, here states a and b. Then ff(a +b) = f(f(a + b)) = f(f(a) + f(b)) = ff(a)) + ff(b)), and equally, by induction over n, f^n (a + b) = f^n (a) + f^n (b). With F_n = f^n (Z) we can immediately conclude that F is linear if and only if f is linear.

[12] In generally intelligible descriptions of the chaos theory one sometimes finds the explanation that the trajectory runs through all points of the space segment that is defined by a strange attractor. This usually is a sufficient description for practical purposes. Mathematically, however, it is not quite correct, because it must be pointed out that, according to the theory of transfinite sets, the state space (as well as any finite segment) is a continuum, which means that it contains a "transnumerable infinite" number of points (power Aleph 1). On the other hand, the trajectory of a real system can only reach a strange attractor of the power Aleph 0, and thus an "enumerable infinite" number of points, and it can do this, of course, only in infinite time. As a matter of fact, in any finite observation time only a finite number of points in the actual segment of the state space can be taken.

trajectories that seem to jump around haphazardly in the state space, and that can take any value within the strange attractor (if one exists). Sometimes it is indeed possible to make a short-range prognosis concerning this sort of system, e.g. from one state of the system to the next. This is especially the case with "deterministic chaos", i.e. deterministic systems that behave chaotically. Long-term prognoses are not possible though - to an observer the changes of state will seem quite arbitrary.

Chaotic systems are often extremely sensitive to a variation of the initial states: Even the slightest deviation from the initial state - e.g. by external disturbances - usually leads to trajectories that differ completely from the trajectory that results from the original initial state. This is the main reason why it is almost impossible to make a prognosis on chaotic systems: even with a meticulous measuring of an initial state one can never be completely safe from small measurement errors. The result is a slightly different determination of the state, which produces a completely different trajectory.

After these rather cursory remarks on some of the most important basic concepts of the theories of dynamic systems, I will now take a closer look at the particular concepts and problems that are important to our subject.

2.2. ADAPTATION AND SELF-ORGANIZATION - AN ETERNAL GOLDEN BRAID

In section 1.3. I drew up a rather unsystematic list of features that might be characteristic for the kinds of systems that the cultural and cognitive sciences are dealing with. Now I would like to define these features in more detail. I will lay particular stress on the analysis of the correlation between the concepts, which seem rather heterogeneous at first sight.

Let us start with the concept of self-organization, which is probably one of the best known and most commonly used concepts in the recent discourse on systems theory. There are, of course, already numerous definitions - and even more informal uses - of this concept. One of the most exact definitions, which also explains the (natural scientific) bottom-up approach, comes from an der Heiden (1992, 72), to whom we can refer here:

> "Die durch die dynamische Abhängigkeit zwischen den Komponenten (Größen, Variablen) eines Systems induzierten Eigenschaften und Strukturen eines Systems und seiner Komponenten nennen wir selbstorganisiert."

(We call those properties and structures of a system self-organizing, which are generated by the dynamic dependency between the elements of the system.)

Obviously, this is a definition - similar to the famous definition of autopoiesis by Maturana (1982) - that (a) starts out from the elements (*Komponenten*) and their interactions (*dynamische Abhängigkeiten*) (b) describes the system behavior (*Eigenschaften und Strukturen*) as a consequence of the interactions between the elements, and (c) determines the repercussions on the elements (*seiner Komponenten*) as a result of (a) and (b).

Krohn and Küppers adopt this understanding of self-organization by succinctly remarking:

> "Rein formal sind alle Systeme selbstorganisierend, bei denen die Ursachen für Systemveränderungen im System selbst liegen und externe Einflüsse auf die Systemdynamik nicht entscheidend sind. Mit anderen Worten, autonome Systeme sind, sofern sie überhaupt organisiert sind, selbstorganisiert." (1992 a, 165)

(In a formal manner all systems are self-organizing if the causes of changes in the system are inside the system and external influences are not decisive for the systems dynamics. In other words, autonomous systems are self-organizing, if they are organized at all.)

Krohn and Küppers know, of course, that their claim has not said much so far. Thus, they point out in addition that the special kind of self-organization - the system-specific rules of interaction - cannot be understood without taking the system's environment into account:

> "Ganz allgemein sind es Strukturen in der Umwelt, die als Randbedingungen im System als Selektionskriterien Bedeutung erlangen, weil sie die systemspezifische Wechselwirkung festlegen. Selbstorganisation ist daher immer an die Existenz einer in diesem Sinne strukturierten Umwelt gebunden." (loc. cit., 167)

(In general structures in the environment are important as boundary conditions and criteria for selection because they determine the interactions specific for the system. Self-organization therefore always needs an environment which is structured in this sense.)

Neither an der Heiden nor Krohn and Küppers clearly explain what is meant by "*Systemveränderung*" (system change), "*Systemdynamik*" (system dynamics), or "*Eigenschaften und Strukturen*" (properties and structures). Using the terminology introduced so far, we might summarize the quoted definitions like this: A system is *self-organizing* if the system-specific rules of interaction generate sequences of states that can be explained only from these rules. There usually are system-environment relations that select specific interactions, which means that

from all the interactions that are possible according to the rules, they choose those interactions that result from particular initial states of the system. As we have seen, this results in specific sequences of states from the system's "space of possibilities". *For the system*, the "boundary conditions" that Krohn and Küppers speak of can only mean that the observed history of the system shows particular initial states - and not others - as a result of the environmental criteria. Once this selection is made, the system dynamics proceed according to its system-specific rules, i.e., its own logic.

We notice immediately that in this description the environment only functions as a "selective trigger," in that it rewards particular initial possibilities and rejects others. We shall see that this view is too simple because it is too one-sided, at least with regard to the systems that interest us here. Nevertheless, it has shown up again and again in recent years how fruitful a stringent analysis of complex systems from the point of view of their self-organization can be. The use of cellular automata for the modeling of self-organization in particular (see below, section 3.1.) could convincingly show how highly complex forms of system behavior can be explained on the basis of self-organization by simple rules of interaction (Gerhard/Schuster 1995; Langton 1992; Boerlijs/Hogeweg 1992). Naturally, a lot more stress has to be laid on the part of the environment than is suggested by the concept of self-organization. The systems mentioned in section 1.3. - and also systems of life, i.e. organisms and/or ecosystems - cannot be understood by their specific form of self-organization alone, but also, and especially, by the way they achieve - and, above all, also change - their self-organization in dependence on their environment. That leads us to the second characteristic feature of our systems, namely the concept of adaptation.

Since complex systems always have to be understood as self-organizing systems, they can naturally not realize their adaptation via direct control from their environment. Adaptive systems can only vary their self-organization, in fact, they can give themselves a rule system that enables them to fulfil particular environmental requirements. Self-organization means that the systems realize their interactions on the basis of intrasystemic rules; adaptation means that the rules of interaction can be varied because of particular environmental requirements until the system reaches sequences of states that are adequate *for the system* with regard to the environmental requirements. Before we look at the relations between self-organization and adaptation more systematically, I would like to illustrate these remarks with the example of a computer program, which Jörn Schmidt and I have developed; it is a combination of a cellular automaton (CA) with a genetic algorithm, of which a detailed description will be given in chapters three and four.

Such combinations of programs are usually called "hybrid systems" (see below, section 3.3.). In this hybrid program the CA is designed to model a differentiating society, namely a class society that develops from an agrarian society. The rules of the CA basically imply that different classes develop step by step from a purely rural population in the following way: from already existing classes, e.g. yeomen, other members of other classes, such as aristocratic big landowners or urban dwellers, develop according to a particular probability. Since it is not possible for a rural worker to rise immediately to be an aristocrat or an urban merchant, this is really an autocatalytic system in the sense that specific classes already have to be in place before others can arise. The cellular automaton *models* therefore and *is* a self-organizing and autocatalytic system.

These processes of differentiation can only take place in the specific environment of a social system. This is comprised of a "natural" environment, which imposes the problem of material reproduction; an "intersocial" environment, which includes other competing societies; and finally, also an inner environment, that of psychic systems, i.e. the particular needs of the population. Therefore, any social system has the following problem: It has to realize its self-organization in such a way that it fulfils the requirements of its diverse environments at the same time and as effectively as possible. In other words, it has to solve a classical optimization problem, which as usual consists of (a) requirements that partly contradict each other, that is (b) underdetermined (i.e. if there is any solution, then there is more then one), and that is (c) not clearly defined, because it is uncertain whether there is an optimal solution at all.

As I pointed out, the concrete development of a self-organizing system depends on its actual initial states and the set of rules according to which the system realizes its interactions. Environmental conditions can, of course, select the system's rules of interaction and, above all, its initial states. On the one hand, however, this is a rather general and abstract description, whereas on the other, it does certainly not sufficiently account for the fact that the system can fulfill its actual environmental requirements even if they remain constant. The reason for this is that in most cases a prognosis of the factual development of a complex self-organizing system from its initial states - even if the rule system is known - is not possible, especially not for the system itself. Therefore it is likely that the selection of rules and initial states by the environment instigates the system to trajectories that can be fatal with regard to long-term environmental requirements. A system that haphazardly persists in the rules of its self-organization because it cannot do anything else must

sooner or later break down because of its growing discrepancy toward its environment. The history of human societies is full of examples of social systems that finally could not cope with their environment because they were not (or not sufficiently) able to change.[13]

Thus, its quality of adaptation obviously must comprise further possibilities for a system, especially its ability to *change its own rules*. This ability was introduced into our program in form of a genetic algorithm (GA), a type of program that will be presented in more detail in section 3.2. Here a short description of the GA will do: It is a program that changes the rules of another program according to specific criteria - the environmental requirements. The one program - in this case the CA - runs according to specific rule systems with the initial state of a purely agrarian population. After a while the whole program discovers that the society's development is not suitable. As a result the GA changes the rules of the CA, which continues to represent the system's social development as it runs several times with its new rules. The states that result thereby are evaluated again with regard to the system's environmental requirements, and if the results are still unsatisfactory the rules of the CA are altered again until they meet the environment's demands. Since this process will be described in detail in the next chapters, these sketches will suffice here.

The program is both a self-organizing and an adaptive system. Its self-organization is accomplished by the rules of the CA, which naturally do not change while the CA runs. Adaptation is achieved by the systematic variation of rules, and this rule changing is itself generated by specific rules, namely those of the GA and the rules of coupling between the two subprograms. With regard to the rules of the CA, these rules are, logically speaking, *meta-rules*, i.e. rules that operate on or above the prior rules. The environmental requirements are preset externally by a user and remain constant until the user changes them. This is not necessarily the case, however, and we will encounter programs that generate their own variable environment. The connection between the first level rules and the meta- or second level rules is not one-sidedly hierarchical in the sense that only the meta-rules can control the first level rules. What we have here is rather an interdependence in the sense

---

[13] Think for example of the dramatic examples of the American Indian cultures, who almost without resistance succumbed to European colonization. Obviously, the Indian cultures were not able to respond adequately to their changing environment - they retained their rules of self-organization for too long. The less well-known example of the Vikings on the island of Greenland shows that a particular society can manage quite well with specific rules of self-organization, but cannot cope with a slight change of its surroundings. The social systems in these examples were obviously not adaptive enough.

that the first level rules can also influence second level rules. This interdependence will be shown in more detail in chapters three and four.

If one takes the logical structure of such a formal system as a model for all self-organizing and adaptive systems, which is possible without reducing their generality, one comes to the following conclusion: Adaptive systems that are also self-organizing must always have two sets of rules, which operate on two different logical levels. As in all self-organizing systems, there must first be rules for the interactions of the elements, which realize the development of the system from particular initial states, often by autocatalytic processes. Second, the system must have specific second level rules that make it possible to systematically vary the first level rules, i.e. the self-organization. The rules of both levels have to be coupled in such a way that they can mutually influence each other. Third, the system must have an "evaluation function" so that it can decide if the states that it reaches are fairly compatible with the environmental requirements or not. If a system meets these three conditions, then it is adaptive and self-organizing, which means that it responds to particular environmental requirements and/or changes in its environment by varying its self-organization.[14]

Once again I would like to emphasize that adaptation must not be confused with external control of the system. Only the system itself controls its development, which means that the sequences of its states are generated exclusively by intra-systemic rules from particular initial states. In this sense, an adaptive system is always self-organizing. Unlike purely physical or chemical systems, however, an adaptive system does not permanently follow the same rules of self-organization, as, for instance, Maturana, Varela, and Uribe (1982) presuppose in their formal model of autopoiesis. The additional meta-rules, including the evaluation function, make it possible for the system to allow for its environment, which results in a variability of its structure, i.e., its rules. This, too, is accomplished by a completely intrasystemic rule system, however.

Of course, such a structure can not only be found in appropriately constructed computer programs. There is a famous model of a dynamic system, which shows these aspects quite exactly, namely Piaget's model of human development (e.g. Piaget 1975). Naturally, Piaget does not use the concepts of self-organization and adaptation, but he does recognize two rule systems which successively and interdependently control the child's cognitive development: assimilation and accommodation. During the phase of assimilation, the mind behaves in a self-organizing way,

[14] The brain seems to be adaptive and self-organizing in this sense, since it has both first level rules (neural interactions) and second level rules (changes of these interactions). The evaluations are carried out by the limbic system (Roth 1996; Edelman 1992).

which means that it integrates different perceptions according to specific rules - Piaget calls them schemata - so that they become processable mental contents. In the accommodation phase the schemata are varied and possibly expanded until the perceptions are processable again - in our terminology: until the environmental requirements are fulfilled again. The parallel to our programs is certainly not accidental, since Piaget is known for striving for logical precision, although we did not have him in mind when we developed our project.

Thanks to the analyses in logical semantics (Tarski 1956), the works of Bateson (see below), and their integration into the communication science by Watzlawick, Beavin, and Jackson (1985), it has long since been accepted knowledge that one must differentiate between distinct logical levels of communication, such as language, mind, and social action. The above illustration of adaptation by a hybrid program "only" shows that and to what extent such considerations can be stated more precisely in form of a hybrid algorithm, and how they can be realized technically. Naturally, some extensions had to be made in comparison with the works mentioned above, which will be dealt with in more detail below.

There should be no doubt that in this sense social systems are, or at least can be, adaptive. That is why for a sociological systems theory there immediately arises an important question from these considerations, concerning empirical research: By which social mechanisms do social systems realize their adaptation? Presently, I cannot give a definite answer to that question. I can only offer some speculative thoughts on that matter in the following chapters. In this respect, the concept of adaptation describes a research program in accordance with Lakatos (1971), as it is characteristic for a general theory of dynamic systems anyway.

Adaptation and self-organization are not in logical contrast with each other, but are instead complementary in the sense that they characterize two different and complementary fundamental features of particular complex systems. Certainly, systems are not equally adaptive and self-organizing at the same time. One can observe phases of proceeding in which the self-organization becomes by far predominant, which means that there is no, or very little, variation of the system's rules. In terms of politics, these would be conservative times. The analyses of such phases prove those theorists right who put the weight of their analysis on self-organization. On the other hand, there are also phases in which adaptation predominates, in which virtually revolutionary developments can be noted, where rules are permanently changed, and new rules are introduced. Therefore, systems like the mind or society cannot be described as self-organizing *or* adaptive, but have to be analyzed with

regard to which of the two basic aspects predominates during the period of observation. Thus it makes sense to regard pure self-organization and total adaptation - understood as the permanent and radical variation of first level rules - as the outer ends of a scale, with the system swinging between them in its environment-orientated and self-controlled dynamics. The result of this process is a measure for the characterization of systems, similar to the one we will get to know for the determination of complexity and the classification of dynamic systems into classes of complexity (see below, sections 2.6., 3.1., and 3.3.). If we look at this matter less formally, adaptation and self-organization are two fundamental aspects of complex systems, which are permanently "braided" with each other in the course of the system evolution. It was the insight into these intertwined connections, which I gained from the analysis of our hybrid programs, that led me to borrow the title for this subchapter from Hofstadter.[15]

2.3. EVOLUTION, LEARNING, AND SELFMODELING: SELF - REFERENTIAL DYNAMICS

After this rather close examination of the concepts of self-organization and adaptation, we may dispense with details regarding the problem of evolution. System dynamics that run according to the principles of adaptation *and* self-organization (regardless of the degree to which they do so) - almost inevitably result in evolution, defined as the development into different forms of complexity. A self-organizing autocatalytic system that is adaptive and subject to variable environmental requirements will always have to use both of these basic mechanisms, by which it continuously comes to new and expanded forms. There is another question concerning social systems that is of a totally different kind: *What* is it that changes in a social or sociocultural evolution, i.e. what corresponds to the genes, phenotypes, and populations in biological evolution? It is true that there are numerous proposals concerning this question (e.g. Giesen 1980), as well as rather trivial borrowings from biology, but as far as I can see there are no models yet that are reflected with regard to systems theory. In chapter five, I will address this point in more detail.

---

[15] As a consequence, the classical dichotomy between "internal" and "external" factors of system evolution becomes irrelevant. The selforganized dynamics is always generated by "internal" factors, namely the rules. Correspondingly, the adaptive rule variation always happens on the basis of "external" requirements - whatever these might be.

It is more complicated to define what "learning" means from a systems theory point of view. At the same time, this is a very important area, since learning is one of the most fundamental parts of adaptation - probably the most evolved. One can, of course, simply say that systems learn or are able to learn. This is what I did in section 1.3., because this claim is acceptable in ordinary language. But if one wishes to use the concept of learning in strictly systems theoretical terminology, it is worthwhile to conform to the logical classification of learning processes developed by Bateson (1972), especially because Bateson employs categorizations that can be used directly for the good of a systems theoretical analysis.

For Bateson it is obvious that all complex systems can "learn." The question is, however, on what logical level that happens. Bateson, oriented to the theory of logical types of Russell and Whitehead, introduces five levels, called learning 0, I, II, III, and IV, for his analysis. In our context, primarily levels I and II are relevant, so I will concentrate on them.

According to Bateson, learning 0 is operative if, e.g. response patterns are principally determined by genetic factors, and experience plays only a very small part in their development. The response of a system learning on level 0 to particular environmental signals is almost always the same if the signals are the same, in other words: there is practically no change in the behavior. Bateson illustrates this claim, e.g. with electronic switchings that never change, or with the logical fiction of a player in the mathematical theory of games explored by Neumann and Morgenstern (see below, chapter 4.), who always employs the same strategy (namely the most optimal one for him) in a particular situation of a game. Learning 0 applies to practically every system that responds systematically to its environment at all.

Learning I, which is more important here, applies if a system varies its behavior when a situation repeatedly reoccurs. This was the case, for example, with Pavlov's famous conditionings. Accordingly, learning I is everything that occurs in connection with instrumental reward and instrumental avoidance. According to Bateson we can say that all the phenomena that were examined in classical behavioristic laboratory experiments refer to learning I; that is where the logical limits of this branch of research lie. At any rate, learning I implies that one or more rules are in effect inside the system, which lead it from its behavior at time 1 to its behavior at time 2. We shall call this a *learning rule*. Learning psychologists and pedagogues usually speak of a learning strategy in this context.

Learning I differs from learning 0 in that an almost constant behavior is varied here with a constant learning rule. Accordingly, the difference

between learning II and learning I is the possibility of varying the learning rule(s), or learning strategies respectively. Bateson offers many examples (derived from experimental behavioral research) of the former. It is sufficient for our purpose to note that learning II obviously corresponds with the "learning of learning", which has served as a primary educational goal in curricula and educational planning for at least twenty years.

Bateson described learning III only vaguely, since he himself concedes that such processes only occur with a radical restructuring of the systems in question - e.g. in psychotherapy, in a religious conversion, and in other similar processes. What this means is that in learning III system elements are varied that had to be held constant in learning II - namely fundamental premises on which the system's identity is based. It is no accident that Bateson, a cultural anthropologist, behavioral researcher, and psychotherapist, explicitly points here to the field of psychotherapy, in which it is the therapist's task to get the patient to revise some of the most fundamental components of his personality. Bateson's examples could be interpreted to suggest that learning III only occurs (if it occurs at all) in borderline situations - namely in situations in which the system is faced with its immediate breakdown.

I will not discuss learning IV, since Bateson himself declares that no single system can achieve that. Bateson, a great universal scientist, was still committed to classical behavioral research in the sense that he formulated his logical classification only in the categories of behavior modifications. In this respect (and only in this respect) he treated systems merely as black boxes. The terminology of systems theory introduced here does not allow that, of course, since every analysis of the system has to begin with the interactions of the system elements, as I have repeatedly mentioned. To use Bateson's classification for the good of a theory of dynamic systems, I would like to analyze Bateson's levels with the use of a formal system, namely an (artificial) neural net:

From a logical point of view, an artificial neural net, which I will merely call "neural net" in the following, is, like any other program, just a complex algorithm. It can be characterized as follows: The elements of the system are (artificial) neurons, which can take particular states, namely numerical values. These values are called activations. A link from neuron i to neuron j has a "weight", which is a real number. The interactions between the neurons result from the fact that the activation value of the neuron i moves via the weighted link to the neuron j where it produces a specific activation value. This is conducted from the neuron j to the neuron k where it is converted into an activation value of k and so on. If we describe the activation value of a neuron i as A_i and the link

weight between the neurons i and j as w_{ij}, and if we assume that the activation value of j can also be incited by various neurons *at the same time*, then the activation value of j usually results from the equation

$$A_j = \Sigma A_i * w_{ij} \quad (2.1)$$

So the equation (2.1.) definitely controls the interactions inside the system. Hence, the activation of j depends on how many neurons have an effect on it, in what states they are, and what the link weights are. The development of the dynamics of this system results from the "external activation" of one or more neurons, i.e., they are externally put to particular activation states, which activate the other neurons until the system reaches an attractor state. Thus, such a system is obviously on learning level 0, since it does respond to environmental signals - external activation - but does not change its behavior. It always responds in the same way to the same signals. The important point for our subject, however, is that now we do not have to speak of the behavior of a black box anymore, but can instead precisely lay down how the state of the system (which is characterized as "behavior on the learning level 0") is realized by local interactions of its elements.

The system reaches learning level I specifically by introducing a "learning rule" (I will skip further technical details here). The function of this rule is the systematic variation of the neural link weights until the system has reached its desired state. The "delta learning rule" for example, a traditional algorithm, is:

$$w_{ij}(t+1) = w_{ij}(t) + (\sigma/n * \delta * o) \quad (2.2)$$

where $w_{ij}(t+1)$ is the link weight from i to j at the time t+1. Accordingly, $w_{ij}(t)$ is the weight at the time t; σ is the so-called rate of learning; n is the number of neurons that have an effect on j; δ is the error, that is, the difference between the real activation of j in the final state and the goal state of j; and o is the real activation of j. To put this all in simpler terms: The delta learning rule changes one weight at one temporal step of the net by taking the weight of the previous time step and adding a product, which results from the error, the number of neurons that have an effect on j, and the real activation value.

Our formal system "learns" in the following way: it begins with a particular initial state and an externally preset goal value for the system (the problem to be solved). The system receives its dynamics according to the interaction rule (2.1.). This rule does not change. What does change are the "relations", i.e. the synaptic links. These are changed according to the actual learning rule on the basis of the system's error measurements until the system has decreased its error below a particular

predetermined threshold. If this happens, the system has "learned" and stops.

The important point is that the system learns by varying its interactions. As I mentioned, the rule (2.1.) remains the same. However, the important parameters of the rule, namely the link weights, do change. The "topology" of the system, i.e. the number of neurons that have an effect on the neuron j, does not change. When a network has learned something, it can "remember" it on specific conditions: if a system has learned to solve a problem after learning to solve a previous one, and it is confronted with the first problem again, it will solve it immediately - it has changed its behavior in the learning process. Obviously, this is what Bateson called learning I.

We should take special notice of the respective learning rule here: Obviously, it is not a rule that controls the interactions between the neurons; that is done by the equation (2.1.). After all, the learning rule is a meta-rule on (2.1.) only in a limited sense, since it does not change the *form* of the rule (2.1.). In that sense, however, it does "control" the rule (2.1.), and is a meta-rule with limited effectiveness in relation to it. Of course, one can conceive meta-rules that change the form of the rules as well, e.g. by turning the summation of (2.1.) into a product. It is normally a question of usability whether this is worthwhile with regard to formal systems. Some of these variations of meta-rules were tested systematically with "genetic programming" (Koza 1992). When real systems are simulated, the question of adequacy, i.e. the empirical question, arises: How radically does the real system vary its rules of interaction?

Such a formal system would reach level II according to Bateson, if it would vary its learning rules, e.g. if it would change from the delta rule into the rule of Hebb, which is equally rich in tradition. This can be achieved by making the net "hybrid", i.e. by coupling up a controlling program to the neural network. It would be the task of this controlling program to decide on the modifications of the respective learning rule on the basis of "meta-learning rules" - e.g. by a genetic algorithm (see below, section 3.2.). Numerous experiments with hybrid neural nets have been developed since (e.g. Belew et al. 1992; Goonatilake/Khebbal 1995), but they have usually dealt with the optimization of the topology - that is, the determining of how many neurons there should be in a network and how many links there should be between them, how many neurons should activate a neuron, and so on. The rules for the variation of these parameters are unquestionably meta-rules, because they can significantly change the system dynamics. Hybrid systems of this kind show learning of the level I, which leads to remarkable performances.

Only recently, it seems, have systematic experiments with learning II occurred (cf. Pearson et al. 1995). At least technically, this can be accomplished without further problems.

Finally, learning III basically entails that the system would give up its "identity" as a neural net and transforms itself into a totally different type of program. We have carried this out, too, at the University of Essen, using a cellular automaton (CA) that transforms itself into a neural net. This net can also revert, becoming a CA again (see below, chapter 4.). Our intention actually was to analyze the self-modeling of formal systems. It seems that two characteristics of our systems meet here: learning III may only be possible in the way that particular self-modelings are performed, which are then modified. This is also suggested by Bateson's references to psychotherapy. It seems that there is still a lot of research and terminological clarification to be done.

So, from a systems theoretical point of view, "learning of a system" is the dynamics of local interactions effected by environmental signals and the reaching of a fixed state (learning 0), the variation of the rules of interaction by meta-rules on the basis of error evaluations (learning I) with new learning taking place on the basis of the already achieved learning successes - the system can partly remember the learned solutions, because it organizes its variation of interactions only with regard to the states it has already achieved; and finally, the variation of the meta-rules (learning rules) again on the basis of the previously realized states and meta-rules (learning II). Even learning III is possible, provided the system has transformation rules that allow it to switch to another system type. With regard to the rules of interaction, these would be meta-meta-meta-rules, since they also vary the meta-meta-rules that make learning II possible.

Compared with formal systems, social systems seem to be able to achieve learning III. We shall return to this claim later. In any case, in the terminology introduced here it is obviously possible and useful to attach a precise meaning to the concept of "learning of systems", which is often used quite vaguely. It is, of course, an empirical question *how* systems realize their actual learning processes.

Since we have already dealt with the concept of self-modeling in some detail in section 1.3., it will suffice here simply to recall this point (including in regard to learning III). I hope that these reflections have made it clear that the dynamics of the systems that are analyzed here can only be understood as "self-referential" dynamics: During its evolution the system perpetually refers to itself: on the level of the interactions (here the individual system states result from each other according to self-organizing rules); on the level of rule variation (here, as in learning I, it is decided which rules, if any, have to be modified on the basis of the

realized states); and on even higher levels, such as those of learning II and III. The interactions that are actually carried out, and states that were really reached thereby, determine whether higher levels of variation are to be called in, and if so, which. This process is accomplished through the evaluation of the system's states according to its environmental requirements.

2.4. THE MEMORY OF SYSTEMS OR: IS IT POSSIBLE TO LEARN FROM HISTORY?

For a system, learning not only means to be able to adapt itself to new problems by varying its rules. It also means the ability to remember successful problem solutions. Every teacher or cognitive psychologist would only recognize learning as successful if an ability acquired thereby could also be applied to a situation *after* the actual learning process. Learning is successful if and only if it includes memory performances. This sounds trivial. But the question of how learning, i.e. adaptive, systems organize memory is not trivial at all. At any rate, models of pedagogy or learning psychology often use a computer metaphor in this context, and speak of a "memory store". Behind this metaphor lies the assumption that the human (and animal) brain stores data, information, and problem solutions just like a computer: A typical program first of all contains the rules (algorithms) by which it operates. Second, it contains a particular amount of data that is entered and with which it has to deal. Third and finally, it contains specific files that store results and perhaps also new information. These files are the program's memory, upon which it can rely at any time. Accordingly, the customary hardware provides special storage chips, whose storage capacity is one of the main components of a computer's power and efficiency.

So far, so simple, in principle. It is certainly no accident that on the one hand the memory problem of computers is usually solved this way, and on the other that this merely technical solution is often thought to be *the* model for memory in general. The kind of storage outlined above consists in a *symbolic coding*, which is composed of bits (a detail that is not so important here, however). In any case, this storage model serves in its basic logic as the understanding of memory both in the ordinary view and in many cognitive psychological theories of memory: a conservation of symbolic codings (e.g. linguistic ones) upon which one may rely when they are required.

There is no doubt that at least several systems actually do with this form of memory. As I mentioned, this holds true for standard computer programs. Obviously, it is partly the case for the mind as well (at least if

it is considered as independent of the brain). And it certainly holds for social systems, which have realized the possibility of introducing symbolic coding in the form of written records as part of their system memory (at least since the invention of writing). But this model obviously does not apply to the brain, which consists only in an enormous number of neurons and their connections, and certainly contains no storage area filled with symbolic codings. From a neurophysiological point of view, "memory is a system property, that is a dynamical feature of populations of neural groups" (Edelman 1992), and in the nervous system there is characteristically no storage in the form of symbolic coding. In the terminology used here this means that memory is a result of specific rule systems, and can therefore be characterized by specific system states. Before we go into this issue in more detail, I would like to draw attention to some problems of the traditional conception of storing.

If memory is conceived as a specific storage component independent of other system operations, the question arises how the use of that store is controlled. In other words: How can it be guaranteed that the contents of the store are called in to serve the system as a memory performance? Logically, the problem can only be solved by introducing specific meta-rules, which halt standard operations of the system according to specific criteria in order to consult the store. As a result, the problem alters its emphasis, first to the criteria themselves (which would finally have to be explained on the basis of the system's evolution), and second to the binding force of the meta-rules: It has to be determined that the system, simply as a computer program, really consults the store in the respective system operations, and that it is not free to decide if it wants to remember or not.

This problem of the binding force is not easy to manage, as will be shown quite clearly with the example of social systems. Even if system memory is only understood in terms of symbolically coded storage, it can obviously not be taken for granted that social systems do remember. The age-old question whether it is possible to learn from history (and if so, *what* can be learned) always depended on the assumption that the memories of one's own past consist in the possibility to consult the symbolically coded contents of the store. For the system, however, it is arbitrary whether this is really done, or whether the history that is remembered in this way is made the basis for new system operations - i.e. actions. One could even assume that there is typically no learning from one's own history this way.

However, it is a well-known fact, not just to neurophysiologists, but also to every learning psychologist and pedagogue, that the systems with which they deal constantly learn from their history and especially from

their learning history - i.e., they perpetually remember their successful learning performances and make those the basis for their further system operations. This is not optional for these systems but is instead an integral part of their system dynamics. According to the traditional storage model this can only be explained (if at all) by use of clumsy *ad hoc* hypotheses. Hence, neither neurophysiologists nor the modern AI seriously follow this model. Let us therefore take a closer look at what could be meant by the concept of memory as a dynamic system property.

The most famous example of a single individual learning process is probably that of Pavlov's dog. This dog is known to have learned to combine two different environmental stimuli - food and bell signal - so that after some time it produced particular behavioral characteristics (namely salivation) even if only the bell rang. So after the learning process, the system "dog", when affected by a particular environmental change, adopted a system state that it did not take in response to the same environmental variation before the learning process. Since the system adopted the same state again at later environmental variations, it also remembered what it had learned.

At least since Hebb's pioneering model (1949) it has been possible to explain this learning process neurophysiologically by the postulation that two groups of neurons, which had no or only an inhibitory connection before, were put in an excitatory connection - namely the group that is stimulated by "food" and the one that is activated by a particular auditory stimulus. "Learning", as generally defined in section 2.3., is the variation of specific rules (in this example the changing of connections). In the language of the artificial neural nets this amounts to the variation of link weights on the basis of a learning rule. The dog's brain varies the connection until a specific mutual state is reached for both groups of neurons. This state is the result that was generated in the brain by the environmental signal "food and bell". At this point the dog's brain (as would be the case for a human brain) does not "know" at what time the groups of neurons should cease changing their state so that a particular state always corresponds to a specific environmental signal and produces an equally definite behavior. Therefore, we can assume that only those states can be regarded as learning results that correspond to simple attractors of the subsystem of the two neural groups - attractors of the period 0 or of small periods (cf. e.g. Hopfield 1982). So learning means that specific attractor states are reached in the individual groups of neurons by the variation of rules.

If the dog is confronted with both signals again after the learning process, they will be received by groups of neurons that have already been varied. It must be mentioned, however, that these groups of neurons

usually will not then be in the same state as they were as a result of the learning process, because in the meantime many other signals had to be processed, and the groups of neurons will have been induced toward other states by external stimuli. At this point the dog has to remember (and this is the only way that the concept makes sense) - the signals (which were the release mechanism for the learning process) now encounter a neural subsystem whose rules (namely its link weights) have already been adapted to these signals. The arousal of the neural subsystem by the signals means that no, or only a little, rule variation has to occur to return the system to the attractor state "activation of the neurons to secrete saliva". Since this process occurs much faster than the laborious variation of rules, an observer can recognize that the system remembers something it has learned by the much higher accuracy and promptness with which it responds to the signals.

So memory, as the "storage" of learning results, does not entail the storage of particular states, e.g. in an external store of the kind mentioned above. It is rather the conservation of particular rules or parameters, like the link weights in our example, that make up the memory. If they remain fixed, or vary only slightly, the system can "remember" in that it can change from a specific initial state, which is generated by the system on the basis of environmental signals - arousal of particular groups of neurons - into an attractor state, which has been realized once or several times before in the learning process. Thus, it *learns* by varying its rules until a particular attractor state can be realized, and it *remembers* by keeping these rules constant, or by varying them only slightly. In other words: If learning ability means adaptivity, and thus variability of the system, then memory consists in restricting this variability.

Pavlov's dog is known for the fact that it also (and primarily) learned to realize the attractor state "salivation" when only the bell rang (i.e. even in the absence of food).[16] So its brain proved capable of interpreting particular environmental signals as *components of a master signal.* In the language of neural nets this means that a particular incomplete pattern could be identified by the system, despite its incompleteness, as the original complete pattern (in this case food and bell) to which it then responded by realizing the corresponding attractor state. So memory is not just the ability to recognize learned patterns from the environment, but also to classify new patterns as similar or dissimilar compared with previously learned patterns.

When the bell signal was played more often to Pavlov's dog without being conjoined with the offering of food, the dog after a time failed to

[16] To be more precise, we should of course say: The neural attractor state that leads to salivation on the behavior level was also realized even when the bell rang without its conjunction with food.

respond according to the learned rule system, i.e. it did no longer realize the attractor state "salivation." The dog had forgotten the learned "meaning" of the bell signal. The systems theoretical explanation is that the learned and remembered rules were overwritten, as it were, by new signals and the learning processes resulting therefrom, so that they did not remain fixed. When the new environmental pattern "bell but no food" occurred for several times, it became the cause for a renewed learning process, which corresponded to a new attractor state. The corresponding rules were fixed again, so that with repetition of the bell signal the system "dog's brain" constantly continued to move toward the new attractor state. At this point the originally learned rule system had been forgotten and the new one was remembered. So what was actually forgotten was nothing but the overwriting of one learned rule system by another.

At the same time, this result enables us to define the "meaning" of an environmental signal for a system: This meaning consists in the realization of an attractor state through application of specific rules to an initial state, which had been generated by an environmental signal in the system. Therefore, there is no meaning of the signal "per se." Instead, the meaning of a signal always depends on previous learning processes involving the signal and the rules resulting from them. After a successful learning process, the bell signal has a different meaning for the dog than before, because different rules cause different attractor states to be realized. In more simplified terms, one could also say that the meaning of an environmental event lies in the rules with which the system generates its dynamics and reaches particular attractor states at the time of the event.[17]

The indefensibility of classical imaging theories has been repeatedly emphasized by constructivist epistemologists (e.g. Schmidt 1987). They argue that cognitive systems do not just image their environment "as it is". With regard to the brain this was especially brought out by Roth (1996). We can summarize this point in slightly more formal terms, and in this context the concept of "image" does not seem to be completely unreasonable: The system makes an "image," not in the common sense, but in an acceptable mathematical sense - that is, a mapping of an environmental event, namely one that realizes a specific state in (and by) the system as a result of an equally specific signal. This state then

[17] The conspicuous similarity of this definition of meaning with the famous definition by C. S. Peirce "for what a thing means is what habit it involves" is of course not incidental, if one considers Peirce's contributions to mathematical logics and the logical orientation of his philosophy. This remains true regardless of the fact that, as a pragmatist, Peirce surely understood his definition in a more behavioristic sense.

changes into an attractor state on the basis of the system rules, which is, as I mentioned above, the meaning of the signal *for the system*. If the system learns, the meaning changes. If it remembers, the once learned meaning is kept constant. If the system forgets, it either realizes attractor states that were generated before specific learning processes, or the system cannot reach a simple attractor at all - it becomes chaotic. This may be the case with certain mental and/or neurophysiological disorders.

I would like to illustrate and systematize these considerations with the example of a very simple system: Given a system S that consists of 6 elements. These elements are connected by some rules of interaction and the totality of rules is R. The initial state Z_1 of S is defined by the fact that it is produced by a particular environmental signal of S (e.g. the activation of neurons). From Z_1 the next state Z_2 is generated by R.

$$Z_1 = o\,o\,o\,o\,o\,o\,R \rightarrow$$
$$Z_2 = o\,o\,o\,x\,x\,o$$

and thus $Z_2 = R(Z_1)$. If Z_1 is altered at this point by a new environmental signal, for example into

$$Z'_1 = o\,o\,x\,o\,o\,o$$

and Z_2 is the attractor state that comprises a successful learning process for S, then there are several possibilities (see also Wuensche 1994):

(a) Z_2 can be achieved from Z'_1 with unchanged R, i.e., Z_1 and Z'_1 have the same attractor. In this case, nothing has to be learned, because R can be kept constant. Accordingly, R remains in memory.[18] So if Z_1 is realized again by the first signal, Z_2 is reached again. Thus, we have $Z_2 = R(Z_1) = R(Z'_1)$. This is indeed possible, because the same final state (or attractor, respectively) can be reached from different states with the same rule system.

(b) Z_1 and Z'_1 are "similar" insofar as a slight change of R is enough to come from Z'_1 to Z_2. If the first signal occurs again (which again results in the realization of Z_1 by S), then with a slightly changed rule system the attractor Z_2 also is reached from Z_1. S does not precisely remember R, but it remembers a similar rule system R′, so that the differences between R and R′ are factually irrelevant for the system dynamics of S. So, S has learned and remembered by "neglecting" the differences between R and R′ (learning by default). Thus, we have $Z_2 = R'(Z'_1) = R(Z_1)$.

(c) Z_1 and Z'_1 are strictly dissimilar, so R has to be considerably changed. If this can only be achieved through a variation of the entire R, the problem of forgetting arises, since S cannot change from Z_1 into Z_2 anymore as a result of a radically changed rule system. R, and with it the

---

[18] In anticipation of section 3.1., one can say that Z_1 and Z'_1 are in the same basin of attraction.

original meaning of the environmental signal, is thus forgotten. In some respect this may occur with learning level III, as presented by Bateson, since the hypothetical change to the system in this case is so radical that fundamental memories are "deleted". If S is so small that learning always entails a variation of its entire rule system, then S cannot remember in the long term. Its ability to learn leads to a loss of memory. As a consequence, S has only a short-term memory, which is permanently deleted in the medium range. If, on the other hand, as much as possible is to be kept in memory, then the system's learning ability, i.e. its variability, has to be radically reduced.

If S is so large that only parts of R have to be varied in a new learning process, S can get a long-term memory by "constructing modularity": This means that only specific parts of R are varied in particular learning processes. These parts can be described as modules of R. One module corresponds to one "class" of learning problems here, for example the learning and recognizing of specific patterns of environmental signals. If a pattern occurs in the environment, S checks whether the pattern can be assigned to a previously learned class, so that, in neglect of the details, the corresponding module can come into play. The state that can be reached by that is an attractor that corresponds to an adequate response of S to the new pattern. If the new pattern cannot be assigned to a previously learned class, a new rule module is varied in such a way that the attractor state that is favorable with regard to the new pattern can be realized. The modules corresponding to the previously learned pattern classes remain unchanged, so that S can keep a long-term memory.[19]

As I already mentioned, the problem with this proposal is the question of which criteria are to be followed to classify the new patterns. If the criteria are too rough, then patterns that are too different will be combined in one class and the corresponding rule module will too often generate the same attractors. Then the long-term memory would only be able to keep very rough and general memories. If the criteria are too fine, too many different rule modules will have to be set up, which would soon exceed the capacity of S and greatly reduce its learning ability. Therefore, S has to find a way to attune the criteria accordingly to fulfil both requirements.

Grossberg has solved the problem of artificial neural networks' short- and long-term memory by introducing a "vigilance parameter" that distinguishes the different pattern classes (Grossberg 1982). Of course,

---

[19] As an example, Montana (1995) combined specific artificial neurons in so-called "functional blocks" when he constructed artificial neural networks. Here the variation of link weights only happens block by block, so that the other functional blocks are kept constant.

this parameter has to be determined experimentally for single nets, the so-called ART-networks (adaptive resonance theory). For natural systems we can assume that the extent of such a parameter evolutionarily originates in S, which means that it itself is part of the evolution of S, or, in the case of biological species, part of the evolution of the species to which S belongs. At the beginning of the evolution, the classification is probably rather rough and becomes finer and finer in its course. Holland's classifier system (Holland et al. 1986) is a similar solution to this problem: rules are changed on the basis of a specific problem situation by a genetic algorithm, and then the rules are stored together with the problem (see below, chapter 4.); here, the system starts with rules that are very general (default rules) and can be applied to very different kinds of problems. In the course of the system's learning history, these rules are more and more refined, until there are a number of specialized rules with relatively fine discriminations among the general default rules. In sections 3.3. and 4.3., we shall see that evolutionary strategies of that kind are a characteristic feature of adaptive systems.

Mathematically, however, there is yet another possibility (cf. Wuensche 1994): Since it is often possible to reach the same attractor from different states, S can also try at a new state Z'_1 to systematically search for those rules that make it possible to reach Z_2 both from Z'_1 and from Z_1. In principle, this is not only possible in the case of (b), where there is a great similarity of the initial states, but also in the case that the initial states differ to a greater extent. The reason for this, as will be shown with the example of cellular automata in section 3.1., is that there always are a lot more possible rules than there are possible states for a system. Therefore, there is, in fact, no small probability of finding common rules for Z_1 and Z'_1. The problem, however, is that up to now no algorithm could be specified that allows the finding of these rules in a relatively short period of time. Precisely because there are so many possible rules, the system cannot simply try out all possible rules in succession; the time that would be needed for such an operation would reach astronomical dimensions. If a simple searching algorithm cannot be found, one has to assume that real systems solve the problem of learning and memory more or less in the way that was demonstrated by Grossberg and his ART networks, namely by constructing modularity and restricting variability.

In broad terms we have outlined a definition of a systemic theory of memory. Let us finally return to the question whether it is possible to learn from history and what can be said about the memory of social systems. As I already mentioned above, social systems learn little or nothing at all from the symbolically coded memory stores, that is, written, pictorial, or other records. We have to differentiate here, of

course: *Per se*, the symbolically coded stores are meaningless, because they are placed outside the system operations, which means that they do not necessarily play a part for the variation and determination of rules. As we have seen, this is different in the case of "storages" that are included in the rules themselves. Therefore, symbolically coded stores are of significance for a social system's memory if and only if the access to the stores is itself a part of the rule system in the form of a rule module or modules. This is for example the case in parts of the social system of science, where the access to symbolically coded previous results is an established part of the methodical repertoire. As with every memory, the reasons are on the one hand economical: The scientific wheel does not have to be re-invented again and again. On the other hand, it concerns questions of priority, problems of plagiarism, and, especially in the humanities, the perpetually revised discursive treatment of historical models. In this respect, the access to symbolically coded stores is a part of the scientific rule system and thus an expansion of the storage capacities. More examples could be added to this one.

The situation is different with other social systems, e.g. with politics or economics. It would be difficult to find the symbolically coded storages by which the operations of the policy system orient themselves. In contrast to science, there is no obligatory knowledge that could serve as the basis of further learning processes, or be the content of obligatory memories. Remembering instead takes place in the form of the rule determination described above, in fact, in the way that environmental signals - problems that are to be solved politically - are classified according to their political pattern or pattern class respectively: foreign vs. domestic policy, allies vs. opponents, own followers vs. those of the political rivals, etc. After the classification into a rule module, this rule module either is slightly varied or it is not, depending on the familiarity of the problem; the criteria of classification themselves may have to be changed, for example, if traditional concepts of an enemy dissolve and simple, bipolar conceptions of the world no longer apply. One could observe this sort of development in great detail at the end of the Cold War. The solutions themselves have to be defined as realized attractor states that are produced by remembered rules. For example, the environmental signal "external threat" is almost automatically responded by an increased readiness of defense, for example the call up of reservists, which is the realization of a new system state. The case is similar with economic systems.

In this context we can refer to an informal solution to the memory problem that was developed by Parsons: In his famous AGIL-scheme he identifies the L-subsystem as the one that is responsible for latent pattern

maintenance. At first, this sounds like hypostatizing specialized stores with symbolic codings. This impression is revised for example in "The American University" (Parsons/Platt 1973), where there is a detailed description of how the university, as an important factor of the society's "fiduciary system", ensures that the students are socialized according to the general social rules by means of college education. So the fiduciary system can only serve its function by both handing down - and thus determining - social rules, and declaring them as obligatory norms of socialization. This guarantees, at least in principle, that the whole system remembers these rules in the sense stated above - for example, by solving domestic problems in basically the same way as in the previous history of the system. An examination of the society of the United States tells us that it is the fiduciary system's scope that restricts and determines how well that society works: The breaking away of some ethnic minorities and economically underprivileged people from the fiduciary system is a threat to the memory of the system, and makes its integration doubtful.

It seems to me that Parsons intended a solution to the memory problem similar to what we formulated more strictly above. In any case, even a cursory look at ethnological and other comparative studies of society tells us that the handing down of determinate rules has always been seen as the memory of social systems. It seems that this has always been a widely held view among social scientists - in contrast to many psychologists, epistemologists, and computer scientists. So in this respect too, general theories of complex adaptive systems can gain a lot from the social science tradition.

Before defining certain positions regarding the theory of science by examples of "complexity", "emergence", and "systems theory thinking", we will now make a little excursion into the theories of decidability and computability. The reason for this is that this study treats theories that are characterized by formal models, especially computer programs. Therefore, it is important to be aware of the fundamental possibilities and limitations of such theoretical and methodical approaches.

2.5. GÖDEL, TURING, AND MÜNCHHAUSEN: THE PARADIGM OF UNIVERSAL COMPUTABILITY

In section 1.2. I explicitly pointed to the central role of the computer in the new systems theories, both methodologically and as a "disciplinary matrix" - systems "are" computers. However, the fascination caused by the computer in this double sense is definitely much older than its technical realization: What lies behind it is the age-old dream of the computability of all problems by means of constructing formal symbolic systems that can precisely depict and solve every problem simply

through formal combinations. As regards the history of ideas, the computer belongs to a tradition that began in Europe with the "ars combinatoria" of Raimundus Lullus. It had its philosophically most important expression in the "mathesis universalis" of Leibniz. And this tradition reached its temporary peak prior to the invention of the computer with the development of mathematical logics in this century.

It is no accident that the search for a formal universal language was always combined with attempts to construct automatic computing machines. In fact, this is logically consistent because the possibility of a universal symbolic system that can depict and solve every problem by computations implies the possibility of transferring these computations to a "computer" - if "computation" means that formal operations are carried out according to definite rules, and that the operations' end means the solution of the problem. Correspondingly, the list of mathematicians and logicians who set about constructing at least virtual computers reaches from Pascal and Leibniz to Turing and von Neumann. It is well known that the history of the "real" computer began with the last two names.

It is not without a certain irony that these last developments and the limits of their approach were demonstrated at the same time. The famous "theorems of limitations" of Gödel, Church, and Turing showed that formal systems - if they have a certain level of complexity at all - are always undecidable, and in most cases even incomplete. This entails the following:

Generally speaking, an algorithm is a computing instruction, that is, a formal operation that makes it possible to work on a particular problem continually in the same way until it is solved. The summation algorithm for the solution of linear equations is such an algorithm: Two equations with two variables each can be summed up in a linear way, that is, they are added or subtracted until only one equation with one variable remains. This procedure always works, regardless of what the equations look like in detail, or of what they might "mean". With regard to formal systems, an algorithm is either a construction rule that allows the forming of new formulas (in a very schematic way, as it were) or, more important here, it is a *procedure of decision* on the question whether a specific formula is or can be correctly derived from axioms.

The possibilities of constructing formal systems of high complexity became crucially important for mathematical pure research: The paradoxes made manifest in the development of Cantor's theory of transfinite sets urgently required a logical foundation for mathematics. At the same time, the possibilities of a formalization of mathematics seemed to open up a way to finally solve all open mathematical problems with the use of appropriate algorithms. It was with this in mind that Hilbert

formulated his well-known "*Entscheidungsproblem*" (problem of decidability): Is there a purely algorithmic way to find out if any mathematical proposition can be proved or not (and if so, how)? It is well known that this *Entscheidungsproblem* must be answered in the negative. There can be no algorithm of decision for formal systems that are as or more extensive than first order predicate calculus, that is, the formalized logics of syllogisms - "All humans are mortal, Socrates is human, therefore Socrates is mortal". This is the "undecidability" of the predicate calculus that was proved by Church. Strictly speaking, this means that there is no general procedure of decision for every predicate calculus formula that allows a decision on whether or not these formulas are derivable from the actual axioms of the predicate calculus. It may be possible for every *single* formula to either derive it, or to prove that its negation is derivable. But there is neither a guarantee that this succeeds with every formula, nor is there a general procedure to carry this task out once and for all.[20]

Gödel actually tightened Church's results with his famous theorem of incompleteness (*Unvollständigkeitssatz*) by showing that formal systems with the extensiveness of number theory (the theory that is necessary for elementary arithmetic) are "incomplete" in principle. This means that there are "true" arithmetical propositions that are not derivable from the established axioms of number theory. The formal system cannot contain all true propositions in principle, even, in fact especially, if they can be expressed in the formal language of the system. Thus, a formally incomplete system provides, as it were, a doubly negative answer to Hilbert's *Entscheidungsproblem*: On the one hand, an incomplete system is always undecidable, that is, there can be no general algorithm of decision. On the other hand, there are "truths" that cannot even be proved in principle, for example by intuition, luck, and/or patiently trying, in the system. The latter could, for example, be taught to a computer program - the classical literature on AI is full of examples of programs that contain heuristic strategies for the solution of every possible problem.

It follows from the theorems of Church, Turing, and Gödel, as results of the new "theories of computability", that there is a class of mathematical problems that cannot be solved analytically. As a consequence, such problems can be processed on a computer only if the program is run without a guarantee that it "stops". You could also say that these programs are not "algorithmically compressible" (cf. e.g.

---

[20] As a "positive addition" to Church's theorem, Gödel proved the so-called completeness of the predicate calculus by showing that every "true" formula is derivable from the axioms on principle. This means that there is a derivation for such a formula. But nothing is said about how such a derivation can be found - by man or machine.

Barrow 1994). We shall come back to this point later (section 2.7.). These are not limitations on the presently available hardware or on the efficiency of the present programs, but are instead limitations on traditional formalization. Chaitin (1987) could show in general that the Gödel-Turing theorems are merely special cases of a very general class of theorems of limitations. Obviously, the limitations of formal systems shown by Church, Gödel, and Turing are a fundamental characteristic of computability in general.

For non-mathematicians, as well as most mathematicians, these theorems are probably not very relevant in practice, because they hardly play a role in normal contexts. For the subject of a systems theory that orients itself to the disciplinary matrix of computer modelings, however, these theorems are quite important for the following reason: If the heuristic orientation to systems that *are* computers (section 1.2.) refers to logically equivalent systems like the formal systems of the type we just examined, then those systems can never acquire the complexity of, for example, the mind. The characterization of our systems in section 1.3. would then mean that they simply are no computers. For the mind *is* able to transcend the limitations of computability, specifically by constructing proofs like those of Gödel, Church, and Turing. Something is obtained by these proofs that the formal system cannot see, namely its own limits - it cannot see what it cannot do anymore, to vary a remark from Luhmann. The mind can, in fact, see this about itself. But if it were merely a formal system, it could not see itself in this way. It might be helpful for a logical analysis of this ability to take a look at what these proofs are actually like:

Without in the least detracting from the achievements of the great mathematicians Gödel, Church, and Turing, one can say that their proofs (particularly those of Gödel and Turing) basically follow a model of proof first realized by Cantor in his classic procedure of diagonal ("*Diagonalverfahren*"). Cantor showed thereby that the set of real numbers had a greater power than that of the rational numbers; in other words: there are different "sizes" in infinity. Basically, the proof runs as follows: Let us assume that we can write down all real numbers as infinite decimal numbers in a particular sequence (if the "size" of the real numbers would be the same as that of the rational numbers, this would be possible "in principle"). We thereby achieve an infinite matrix whose lines consist of the single real numbers, that is, one infinite decimal number each, for example 0.33333 ...; and whose columns are the corresponding digits of a decimal number at their respective places. Now we form a new number by the following procedure: as the first term, we take the first term of the first number in our matrix, subtracted by 1; as

the second term, we take the second term of the second number, again subtracted by 1; the third term is of course the third term of the third number, again subtracted by 1, and so forth. Since the construction of our new number accesses the diagonal elements of the infinite matrix, this procedure is called procedure of diagonal (*Diagonalverfahren*).

It can be shown that this new number is not yet included in the matrix - in contrast to the assumption. Since this procedure can be endlessly repeated - the application of the same procedure after the new number is added to the matrix again results in a new number - it follows that every matrix of this kind is incomplete, and thus that the real numbers have a power greater than that of the rational ones (which, in fact, are infinite as well).

The philosophical problem underlying such a treatment of infinity cannot be discussed here. My concern is the "idea of proof", which basically consists in the following: A formal system - in our example that of the real numbers - is observed *as a whole*, and thus "from outside." In particular, this view assumes that the system is "complete". Afterwards, a new element is introduced into the system that is made up of parts of elements that were already there. At the same time, the parts are varied. So actually, by expanding the system one comes back to it. The proof that this expansion is a real one is the proof for the fact that the system is in some respects incomplete.

In this way the incompleteness of the number theory (Gödel's theorem) can be proved by the construction of new number theory propositions on this principle (cf. the description in Hofstadter 1985). Accordingly, this can be done with the undecidability of the halting problem by showing that the system of computable functions is incomplete in this sense. The decisive point is, of course, the very possibility to departing from the system, that is, to reach a meta-level in relation to the system. Returning to the system is then the argumentational "trick" (although not a "cheap" one) that provides the result that had to be proved. This procedure, which transcends the system's possibilities, can be described with an apt expression by Tipler (1994) as the "outgödeling" of the system: doing something that the system simply cannot do.

Several critics of AI concluded from these results that no computer could ever be "intelligent", because it cannot "outgödel" itself (cf. e.g. Penrose loc. cit.). However, this argument is cogent only if computer programs are seen as formal system of the kind we just examined, that is, as systems that cannot go out of themselves. In sections 1.3., 2.2., and 2.3. I already mentioned that there are programs that operate on different logical levels, and that this is indeed necessary for the adequate recording

of systems as the society, mind, or language. We can now state this point more precisely.

Naturally, every computer program continues to be a formal system, since it runs according to formal settings - algorithms. In particular, the program continually carries out all kinds of computations. So, inasmuch as real systems are modeled as computers, the "paradigm of universal computability" remains valid. However, the meaning of the term "formal system" has to be expanded so that the fundamental possibility of outgödeling itself can and must be implemented in such formal systems - however limited this might be. In other words: The system must (a) consist of at least two logical levels - often implemented by different subprograms; (b) be able to "go" from one level to the next and back; (c) be able to model itself on one level, that is, it has to be able on a meta-level to make a "picture" of itself on the first level. Once again, we come across the fundamental importance of the capacity to self-model here, which was already emphasized in section 1.3. The system does not, however, necessarily have to be able to make a picture of itself on all levels at the same time, because this would certainly lead to an infinite logical regress.

Basically, programs that consist of several logical levels are formal systems, too, of course, and therefore the theorems of limitations are equally true for the programs *as a whole*. There is no algorithm of decision for the programs, and they are always incomplete with regard to particular propositions, no matter how many logical levels there are. However, this probably goes for all real systems. These programs are basically able to proceed in the way the mind does when it is outgödeling single systems: The program can transcend *every single system* it consists of, and so it can *de facto* check every single problem for its solubility. What becomes crucial here is the fact that, on the one hand, programs are formal systems, but on the other they are physically real as well - in fact, they operate in real time. To borrow from Luhmann again: The program practically circumvents the limits of undecidability and incompleteness by transforming the actual problem into the temporal dimension, that is, it has a conception of succession.

Thus, there can be operational instructions that basically imply the following: "If you get stuck on one logical level, then go one higher, analyze the lower level from there and expand or change it, until you come to a solution on the first level." So the "outgödeling" is achieved by the fact that the program is a *dynamic* system that is not invariably given like abstract logical and mathematical systems. There can also be operational instructions of the following kind: "If you get stuck, construct a higher level *yourself*, from which you can analyze the first level." There

should be no doubt that the mind, too, can achieve its outgödeling performances only by the fact that it operates in the physically real time, that is, in succession.

I think it is time for a brief semantic remark here: The meaning of the "different levels", in which one lies "above" the other, is of course a merely metaphorical one. There "are" no levels, neither inside the program, nor in the brain. Again, the *logical* difference between the single levels is realized by succession, namely by different loops that the program runs in appropriate sequences. The logical step from one level to a meta-level always means that the program stops its operations relating to one loop according to specific criteria, and goes over to another loop. The results of the operations that are carried out there, as long as they refer to the first loop, are transferred to that loop and assume the character of rule alterations and other new instructions. Therefore, the spatial level metaphor is just a metaphor that orients itself to the terminology of logical semantics, from which it derives its precision. No real system has levels, however, but "only" the succession of different loops.

In some respect, this situation certainly bears a strong resemblance to the famous story of the Baron of Münchhausen, who pulled himself and his horse out of a swamp by his hair. In contrast to that story, however, we are not dealing here with a cheap trick, but with logically precise algorithms which allow an exact analysis of the development of order and of complex systems' action competence. This possibility has often been examined in a general form (see e.g. Dupuy/Varela 1992); apparently, that concept can be made fruitful both precisely and quite practically with programs of the kind sketched out above. Thus, the connection that was made in this subchapter's heading between the great mathematicians Gödel and Turing and the lying baron is more than just an aphorism: It serves to demonstrate that the idea behind the (physically quite improbable) story of Münchhausen is obviously becoming quite relevant.[21]

---

[21] The equally well-known story about the frog who had fallen into a glass of milk and was in danger of drowning refers to the fact that on particular conditions even Münchhausen's story might be physically plausible: Through its own activity (constant wriggling, as it were) the frog caused the milk to curdle into butter, after which it could leave the barrel. So, the frog performed something like pulling itself out of the barrel by its own hair, if you like.

2.6. COMPLEXITY AND EMERGENCE: CONCEPTS AND THEIR VAGUENESS

There are presumably no other scientific concepts that have been used as frequently over the last years as those of complexity and emergence. At the same time, no other concepts have been used as indistinctly and ambiguously - at least in the mathematical sciences, which strive for exactness and logical precision. Somehow, scientists seems to believe they have some understanding of these concepts, and somehow they also succeed in forming meaningful sentences with them - I have already made such an effort many times myself. But it is a totally different question whether these concepts can also be defined in a way that allows a useful working with them.

In a attempt at systematization, Lloyd has compiled a list of definitions of the concept of complexity, encountering no fewer than 31 different definitions. In the following, I will quote an excerpt from this list to give an impression of its conceptual heterogeneity (after Horgan 1995, 61).

Complexity is

- "Entropy: the complexity of a system is equal to its thermodynamical measure for its actual disorder, that is its entropy.
- Information: complexity is proportional to the ability of a system to give information to an observer.
- Fractal dimension: the fuzziness of a system, measured as the degree of details which it shows in arbitrarily small scale.
- Effective complexity: the degree of regularity of a system.
- Hierarchical complexity: the diversity of different levels of a hierarchically structured system.
- Grammatical complexity: the degree of universality of a language which describes the system.
- Thermodynamical depth: the number of thermodynamical resources required to generate the system.
- Temporal computational complexity: the time a computer needs to describe or compute the system.
- Spatial computational complexity: the store capacity which is needed by a computer to compute the system.
- Informational distance: the degree to which a part of the system differs from others."

A more heterogeneous list is scarcely conceivable, and it must be added that other suggestions, especially from the theory of computability (such as, for example the concept of the "logical depth" of a system) (Bennett 1988), are not even listed here. Horgan (1995) uses this heterogeneity as an opportunity to doubt the meaningfulness of the whole theory of

complexity. But an overly simplistic view can be deceptive here too - one can show that this heterogeneity is smaller than expected, and observe how a partial heterogeneity can prove quite useful: There is a definition of complexity from classical cybernetics that still practically underlies most of the newer ones, namely the determining of complexity as the *number of possible states* of a system (e.g. Beer 1974). If we abstract for the moment from the problem of determining this number for a system, we can easily see that some of the above definitions can be reduced to that standard determining.

Let us begin with the concept of entropy. From the time of classical thermodynamics on, entropy has been defined as a system's degree of order *at a particular time* - the famous second fundamental theorem of thermodynamics is known to postulate the increase of entropy in closed systems. Therefore, a system's entropy results from the fact that the state of the system is more or less probable, that is, more or less distant from the state of total disorder ("equal distribution"). In textbooks, this is usually illustrated by the example of a gas in a locked container, whose molecules are concentrated at one point in the initial state - the "improbable state" - and then spread out more and more evenly over the whole container in the course of the system time - the reaching of "probable states". Thus, if one wants to define the complexity of a system with regard to the concept of entropy, one has to ascertain the degree of order or disorder, respectively, that is possible for the system. In other words: in this definition, the complexity of the system results from the question of which states can be taken by the system at all, and which entropy has to be assigned to the single states. It is of special interest here, of course, to what degree the actual system trajectory shows an increase or decrease of entropy, or an alteration of its increases and decreases. Thus, this definition of entropy is just an extension of the above mentioned standard definition of complexity, where the number of possible states is combined with their corresponding entropy.[22]

The presentation of complexity with the concept of information is an application of the classical theory of information by Shannon and Weaver (1963), where information results mathematically from negated entropy. Thus, this definition is not strictly different from that of entropy - it is just a simple mathematical variation.

The two definitions of computational complexity are simplifications of the concept of "algorithmic complexity" by Chaitin (1987) and

---

[22] This is shown very clearly, for example, in Langton (1992, 60), where complexity is defined after Shannon's theory of information as $\Sigma p_i * ld\ p_i$, according to which p_i comprises the probability of the possible states, and ld is the dual logarithm. In sum, this results in the entropy of the states (their content of information, to be more precise).

Kolmogorov. It states that the algorithmic complexity of a system equals the length of the algorithm a computer - or a human being - needs to compute the system behavior. (It is not important here whether the length is determined in the form of lines of a computer program or in bits.) A system whose dynamics does not follow a discernible regularity has an algorithmic complexity that is, in principle, as huge as the number of states that the system can take: It can only be described by successively indicating the different states. For example, a sequence of numbers that consists solely of random numbers can only be described by printing its exact sequence. On the other hand, a sequence of numbers that only consists of the repetition of one number - like 1,1,1,1,1, ... - can be described by a very short algorithm, for example by "Keep on printing one '1' after the other." If a system's algorithmic complexity is smaller than the length of the sequence of states it is termed *algorithmically compressible*. This means that there is an algorithm whose length is smaller than the number of states that (can) appear in the course of the system dynamics. Thus, in contrast to random sequences, algorithmically compressible systems are determined by discernible regularities. They have, as Gell-Mann calls it, an *effective complexity*: we define the effective complexity of a system as the length of the schema by which the system's regularities can be described (Gell-Mann 1994); for the concept of the schema see also Gell-Mann 1992).[23]

Respectively, the sequence 1,1,1,1, ... has a very low effective complexity, because its *regularity*, as well as the algorithm that determines its proceeding, can be described by a short representation. Thus, both effective and algorithmic complexity are very low here. A purely random sequence, on the other hand, has a very high algorithmic complexity, but a very low effective complexity, namely precisely zero. As a result, we can state that there is a covariation between algorithmic and effective complexity.

Accordingly, a "schema" is an algorithmic compression of the system dynamics: It is easy to see that its length depends on the number of possible states, inasmuch as a very small number of possible states always implies a short schema - all you have to do is to name the states and their (possibly statistical) frequency. This would be the longest schema for this system. Algorithmic complexity increases with the

[23] Gell-Mann does not say whether his introduction of the concept of the "schema" was inspired by Piaget or Piaget experts. However, there is a striking similarity to Piaget's corresponding terminology, especially in Gell-Mann 1992. By the way, these definitions give an answer to Wittgenstein's famous question of what it means to follow a rule: a system follows one or more rules if there is a schema that algorithmically compresses the system's sequence of states.

number of possible states. In parallel fashion, effective complexity, which is to say, the length of the schemata needed, increases. Of course, this does not occur in a linear manner (or does so only in exceptional cases). Both algorithmic and effective complexity can fluctuate greatly within a given number of possible states. In the last instance, however, the degree of algorithmic complexity depends on how many states are possible: If there are very many possible states and a low effective complexity, the system's algorithmic complexity is very high.

On the whole, algorithmic complexity comprises a more precise elaboration of complexity's standard definition. By itself this elaboration obviously would be insufficient (since, for example, chaotic systems would then have the highest complexity), and a useful degree of complexity regarding content (that is, one that corresponds to the intuitive ideas of complexity to some extent) can only be given in combination with effective complexity. It is quite clear, however, that the definitions that we have looked at so far are not that heterogeneous at all. They can instead be regarded as variations and expansions of the standard definition.

As a matter of fact, grammatical complexity depends on the question of how the concept of system complexity is defined. Viewed apart, a descriptive language has to be extensive enough to include all phenomena that are of interest. If the latter are not to be determined completely *ad hoc*, the determination has to be accomplished by means of defining the composition components of the system, and, as a consequence, which definition of complexity is to be used as far as the complexity of the system is concerned. By this process, however, the definition is again reduced to one of the above - or possibly to another one that could be chosen. "Language" and "grammar" mean a formal system into which the system that is to be described is mapped. To achieve such a mapping, the formal system has to be as "complex" as the mapped one - with regard to the definitions we had so far this means: corresponding to one of the variants of the standard definition. Thus, we do not have an autonomous definition here, but rather a technical possibility of gaining a grasp on the mapped system's complexity through analysis of the formal system.

Of course, the definitions of "fractal dimensions", "hierarchical complexity", and "thermodynamic depth" can not simply be applied to the standard definition. They have a greater autonomy, although the concept of the "number of possible states" is important here too. I will not go into this issue in detail here. Undoubtedly, however (and this is the second aspect of this problem), these definitions (except for that of algorithmic complexity) cannot simply be reduced to the standard definition. They express different aspects of what we regard as complex

with regard to the system *as a whole*. Therefore, the thing to do would be to combine several of these definitions in a non-reductionist way.

Normally, the question of what is specifically interesting about a system depends on the concrete research interest involved. One might be particularly interested in the possible states within the foreseeable system time, and thus in its algorithmic complexity. One might also be interested in the most concise theory (schema) for understanding the system's regularities, and thus its effective complexity - for example: Is Marx' (rather lengthy) theory of historical materialism "compressible" into a sociological systems theory? Within the framework of a differentiation theory of social systems, the distance of information is presumably the primary question, and so on. Such "aspects" of complexity can certainly not be reduced to one another, although there probably is a covariation between them. As a consequence, one should define a total degree of complexity, which comprises all logically independent degrees of complexity, and allows a comparison of systems with regard to their total complexity. If we assume that several of the singular definitions above cannot be reduced to one another, and that there are possibly even more of them, and if we further assume that on the whole there would be a list of n degrees of complexity, then the different degrees of complexity K_i result in an n-dimensional vector. If we call the total degree of complexity K, we have

$$K = (K_1, K_2, \ldots K_n)$$

where the size of K, as it is often the case in vector algebra, is a scalar product of the components K_i, thus

$$K = \Pi_i K_i.^{24}$$

The differences of the total complexity of systems can be compared in this simple way, regardless of the still existing possibility to relate systems with regard to their "singular complexities". You could also say that the total degree of complexity defines an n-dimensional *complexity space*, where a system is plotted as an n-dimensional vector according to its specific total degree of complexity.

---

[24] It seems plausible to define K as the length of the vector, thus $K = \sqrt{\Sigma_i K_i^2}$. The disadvantage of such a definition, however, is that a purely random sequence would have a very high complexity, possibly the highest, compared with other systems. That is why I chose the above definition, according to which a purely random sequence, e.g. the random sequence of a chaotic system, obviously has the complexity K = 0, because the effective complexity $K_e = 0$ in this case. This will be illustrated in more detail with examples of the so-called Wolfram classes of cellular automata in section 3.1.

This results in another prospect of analysis: The definitions of single degrees of complexity we have dealt with so far refer to systems whose rules do not change (as is typical in mathematics and physics), such as sequences of numbers. Of course, the single degrees of complexity K_i (and thus also the total degree K) depends on the specific rules of the system. The concept of effective complexity clearly expresses this dependence. Let us now look at those systems that interest us here, that is, systems that are able to change, expand, and possibly also reduce their rules. It should be obvious that if these rules are changed, the respective K_i (and thus also K) changes too. For example, a chaotic sequence of states can be put in "order" by introducing a rule that periodically repeats specific states. Obviously, this would increase the degree of effective complexity involved. Therefore, K can be determined in such systems only *with regard to an actually valid rule system*, and thus only for the time in which the system does not change its rules. Strictly speaking, K should be given a time index for the time of rule constancy. If the system is observed for a longer period of time, variations of K arise, and thus, with regard to the complexity space, *trajectories* of the system in the complexity space ensue. Therefore, there are always two trajectories that are important for the dynamics of such a system, namely that in the state space and that in the complexity space. This is important primarily because the trajectory in the state space is not very meaningful by itself until one knows which rules (and rule variations) have been in force in order to bring about this trajectory. Quite another possibility, by the way, is to define complexity as the degree to which a system is able to change its rules, that is by its adaptive capacity or its variability.

I will return to these considerations, which have not yet been worked out technically, in section 3.3., and so the preceding remarks will have to suffice for now. First of all, I wanted to show that it is indeed possible to assign a precise meaning to the often rather indistinctly applied concept of complexity, particularly insofar as one wishes to include different characteristics of a system into a *single* concept of complexity. As regards research, it might be enough in most cases to refer to a single or a few aspects of the degree of complexity. However, it is always important to keep in mind that this sort of terminological decision involves a reduction that can only be justified by concrete practical considerations, and that as a consequence possibly does not take other important aspects of the situation at hand into account.

It follows from what we have said so far that a specific task of systematic complexity research is to make a list of the distinct individual components of the degree of complexity that are both indispensable and irreducible to each other (given our present level of knowledge). Once this list is compiled, it will probably have to be perpetually extended. The

very fact that interdisciplinary comparisons show that the various disciplines often use rather indistinct and, most of all, rather different concepts of complexity demonstrates the extraordinary importance of this task.[25] Furthermore, the technical aspects concerning the total degree K must be settled: The single degrees in question (e.g. algorithmic complexity and the distance of information) always have different magnitudes, so that justifiable standardizations have to be made. However, these are tasks for the near future.

It cannot be denied at the end of this discussion of complexity that Gell-Mann also considers the possibility of changing complexity. To this end he introduces the concept of *potential complexity*: to define this one has to look at the possible future histories of the system, i.e. state successions and compute the mean values of the effective complexity of the system at each time. The resulting potential complexity tells us something about the probability of the evolution of the system to a very complex one. (1994).

This definition is influenced to a considerable degree by the quantum theory, which I will not comment on here. Regardless of its theoretical elegance, I think that there is a technically simpler way to manage the procedure I suggested above, even if it appears even more complicated than Gell-Mann's definition. Theorists of complexity may (in fact, *will*) have a tough time with this definition too.

The task of ordering the confusion of the concept of complexity is already difficult, but it is even harder when coupled with the concept of emergence. The main reason for this is that two usually distinct problems come together here: the problem of complete novelty and unpredictability, and the problem of the difference between a system and its elements. This problem conjunction is expressed quite clearly in the following attempts to determine the concept of emergence.

On the one hand, emergence points "to the fact that in the course of evolution new things and events occur, with unexpected and indeed unpredictable properties" (Popper/Eccles 1977, 22). But on the other, "Emergenz bezeichnet das plötzliche Auftreten einer neuen Qualität, die jeweils nicht erklärt werden kann durch Eigenschaften oder Relationen

---

[25] In linguistics, for example, there are elaborate studies concerning the different complexities of languages that fail to offer a precise definition of the concept of complexity (e.g. Comrie 1992, who explicitly admits this). Similarly, Andersen (1992) reports studies about language acquisition which take into account increases of complexity, in which the respective concepts of complexity remain rather vague. Yoffee (1994) rather ironically notices a similar deficiency for the fields of archeology and early history. The situation is quite similar in both the humanities and the social sciences in general.

der beteiligten Elemente, sondern durch eine jeweils besondere *selbstorganisierende Prozeßdynamik*." (Krohn/Küppers 1992 a, 7f., original italics).[26] The first quotation concerns the occurrence of something new. The second concerns the same matter, but here the emphasis lies on the fact that this novelty cannot be explained by the elements and their relations, but instead (as one may paraphrase Krohn and Küppers), only from the system's behavior as a whole.

Obviously, there is a certain lack of clarity here. The succinct summation by Popper and Eccles is, of course, much too broad, because they too would probably not describe every unpredictable occurrence (such as, for example, a lottery win) as emergent. Otherwise, every distinct component of a random sequence would be emergent, and in the end, the concept of emergence would be synonymous with that of chance, or else perhaps with that of chaos. Krohn and Küppers state their position more precisely by defining novelty as the irreducibility of emergent system behavior to its elements. We may assume that they would interpret Popper's concept of "unpredictable" as entailing that an emergent system behavior cannot be predicted if only the system's elements and their properties and relations are analyzed.

However, this still does not clarify things entirely. What is meant by emergence as a "new quality" (*neue Qualität*)? A system always behaves - it has a self-organizing process dynamics (*selbstorganisierende Prozeßdynamik*) - in the sense that its elements interact with each other, which generates different states. If these interactions are known precisely, then one can also explain the individual states, and thus the dynamics of the system. The required analysis is, of course, often quite difficult, and in the case of chaotic systems one soon comes up against the limits of exact prognosis, but Krohn and Küppers speak of an inexplicability *in principle* (and do not confine their definition to chaotic systems).

This can be explained quite well with the example of the previously analyzed fox and goose system: I mentioned earlier that at specific parameter values spiral patterns arise from this system, which entail a "dynamic stability" for the system. This system dynamics occur, for example, if the reproduction rates of foxes and geese are set in the ratio of 2:3. Thus, the "explanation" of the system behavior can be drawn from the observation, or determination, of specific relations - or rules of interaction, to be more precise - between the elements, from which the pattern formation evidently ensues. Of course, this ratio of reproduction rates is not a property of the foxes or geese, but is a relation between

[26] Emergence is the appearance of a new quality which cannot be explained by properties or relations of the participating elements but only by a particular self-organizing dynamics of processes.

them. Thus, it is insufficient for an explanation of the system dynamics according to Krohn and Küppers. Consequently, the spiral patterns are not an emergent quality, but are something else - but what?

The problem with the concept of emergence is probably that it is usually hopelessly overloaded. Cariani also notices this when he critically remarks in a systematic analysis of emergence: "The problem of emergence ... in its most general form encompasses all questions of fundamental novelty in the world, the origins of order, and the increase in hierarchical complexity" (1992, 775). A concept that is meant to include all problems of the world at once can, of course, not be defined precisely and generally at the same time.

Therefore, Cariani introduces a distinction between three senses of emergence. He designates these senses "computational emergence", "thermodynamic emergence" and "emergence relative to a model". The third sense, which is of special interest here, is characterized as follows: "Hence while computational emergence is concerned with what kinds of global behavior can be built up through the operations of local rules without altering the local rules, emergence relative to a model is concerned with the formation of global structures which then constrain and alter local interactions" (loc. cit., 778). And he continues: "The emergence-relative-to-a-model-view sees emergence as the deviation of the behavior of a physical system from an observer's model of it" (loc. cit., 779). From these definitions Cariani draws the conclusion that it on the one hand makes sense to speak of emergence only in the emergence-relative-to-a-model-view, because the formation of something new (and unexpected) can only be expressed with regard to the present knowledge. Yet on the other hand, one cannot speak of this kind of emergence in the case of computer programs, because they can only produce a "computational emergence". It is presumably already clear that this is not necessarily the case. But let us first look at what part a theoretical model plays.

I think that if one wants to include both of the above mentioned considerations into the concept of emergence (as almost all writers on this problem do), one must always apply emergence to a theory or model of the systems that one wishes to analyze with respect to emergent phenomena. Then, of course, the concept of emergence loses all its mysteriousness. When we speak of emergence with regard to a system, this only means that our model of the local interactions cannot (or cannot sufficiently) explain the emergent global system behavior. In other words: our theory (our model) of the system, in the terminology of local interactions of the elements, is undecidable, or even incomplete, with regard to the emergent phenomena (in the logical sense explained in

section 2.5). As a result, there are "true" propositions, namely concerning the observed emergent phenomena, which are either not derivable from the theory in principle (the incompleteness of the theory), or for which there is no general procedure of derivation from the premises of the theory. As a consequence, the theory simply has to be completed - however laborious or even unachievable such a task may be at a certain time. Thus, emergence is far from being a property of particular systems - a "new quality". Likewise, it is not something that can be defined by specific system processes. It results from an incompleteness of theoretical models, and only because of that it is "new" (*neu*), "unpredictable" (*unvorhersehbar*), and "inexplicable from the relations between the elements" (*nicht erklärbar durch die Relationen zwischen den Elementen*).

Let us explore this view with some well-known examples. The remark that the mind is emergent with regard to the operations of the brain obviously means that the theories of the local interactions between the neurons do not yet allow expression and explanation of the mind - the theories' grammatical complexity is insufficient. Roth (1996) has outlined what such a "systemic" theory of the brain could look like. But as far as I know, this is presently just a perspective, and perhaps a concrete utopia. Nevertheless, there are presumably few neurophysiologists who doubt that it is reasonable, and that it is the great goal of their science to search for such a theory. A systems theory of the brain will especially have to inquire as to what extent not only local interactions produce global states of consciousness, but also how these global states change local interactions. The learning rules for artificial neural nets (see above, section 2.3.) give an impression of what this could look like.

Correspondingly, the same can be said for the actual subject of this study, social systems. A complete theory of society would be one that could explain the occurrence of an "emergent quality" (social systems) from local interactions of single "psychic systems" (to use Luhmann's concept here), or from single social actors, respectively. Here, such a theory also does not exist, and the way to get there is certainly as laborious as that which leads to a systems theory of the brain. There are approaches, however, such as the one developed here, and such a theory will also have to show that, and to what extent, social systems can only be understood as the permanent interdependence ("dialectics", as Berger and Luckmann (1966) call it) of local interactions and the global states that are generated by them, and that likewise vary the system's interaction rules. As a consequence, the meaning of "emergence" in these and other examples depends on the question as to what special way these incomplete theories will have to be supplemented: Which mechanisms (from which the emergent phenomena were generated) can be recognized

for the corresponding system types - mechanisms as they appear in Eigen's hyper cycle (that is, autocatalytic mechanisms), mechanisms of rule variation by learning, and so on.

A relatively concrete example for this view of emergence can be taken from Pimm (1991, 173): Pimm had constructed computer models together with Post to check ecological systems for their "capacity to include". That means that Pimm and Post successively constructed artificial ecosystems from several species, allowed them to develop, and added new species after specified periods of time. (The models were based on equations of the previously mentioned Lotka-Volterra type.) The results were always alike: the ecosystems "settled down" to a total number of 10-15 species, *regardless of the peculiarities of the newly introduced species*. If more species were added, they either died out or displaced a corresponding number of older species. Investigations of real ecosystems confirmed these model results. Lewin, who describes these results quite clearly, calls them a prime example of emergence, but not mystical yet as a product of local interactions (1992).

If computer models are taken as models for "real" systems, they are as complete or incomplete as the theories on whose basis they were constructed - they are actually just the transformation of the theory into a formal model. If emergent phenomena in the form of particular states of a formal system are generated, then these phenomena are in fact not emergent, since they evidently follow from the theory. They are rather (subjectively) surprising, like the spiral patterns in our fox and goose system. However, if computer programs are taken as systems *sui generis*, the repeatedly mentioned systems that are able to vary and expand their rules meet Cariani's criterion: On the one hand, the programs are their own models, and on the other, they are able to realize states that are emergent with regard to themselves by changing their own rules by themselves. Of course, this is only true if the systems are not observed *sub specie aeternitatis*, that is, from the aspect of what might become of them in principle in an indefinite period of time, but *in relation to a particular period of time*. This has already been stressed with the concept of "outgödeling": Programs operate in physical time and so they can only be observed for particular periods of time. Therefore, the concept of emergence, similar to the concept of complexity, always refers to a concrete period. In the same way that the (total) complexity of a system can change in time, the question of what is emergent with regard to the system depends on when this question is put: Phenomena that are emergent with respect to the system immediately lose this quality as soon as the corresponding theories are completed or the computer programs have developed accordingly. For example, the history of the search for a

unified physical theory can be interpreted in the same way: Phenomena that for a long time seemed to be emergent do not seem so anymore (cf. e.g. Peat 1988; Barrow 1994).

To sum up, one can say: In this general definition emergence consists in the occurrence of particular global system processes and states that cannot be explained (yet) from the rules of interaction of their elements, because of the incompleteness of the theoretical models. Any definition of emergence that postulates this as a *fundamental* explanatory impossibility has to bear the burden of proof. On the other hand, the definition that I suggest here can, of course, not prove for any individual case that in the foreseeable future there will be a complete theoretical model for the phenomena described as emergent. It just says that in the absence of counterevidence (which can practically never be produced), one should not attempt to discover emergent phenomena in mystical aspects of systems such as "new qualities", but instead through the solid endeavor to complete the systems' models.[27]

2.7. SYSTEMIC THINKING: THE KANTIAN STANCE AND FUNCTIONALISM

> "In contrast to this, cybernetic explanation is always negative. We consider what alternative possibilities could conceivably have occurred and then ask why many of the alternatives were not followed, so that the particular event was one of those few which could, in fact, occur. [...] In cybernetic language, the course of events is said to be subject to restraints, and it is assumed that, apart from such restraints, the pathways of change would be governed only by equality of probability. In fact, the "restraints" upon which cybernetic explanation depends can in all cases be regarded as factors which determine inequality of probability. If we find a monkey striking a typewriter apparently at random but in fact writing meaningful prose, we shall look for restraints, either inside the monkey or inside the typewriter. Perhaps the monkey could not strike inappropriate letters; perhaps the type bars could not move if improperly struck; perhaps incorrect letters could not survive on the paper. Somewhere there must have been a circuit which could identify error and eliminate it." (Bateson 1972, 399 f.)

I chose this rather extended quotation because I think that Bateson expresses in a very clear-sighted way exactly what is, or at least should be, the modern systems theories' specific way of thinking. Of course, Bateson does not think that there are no causal processes, or that such processes cannot be comprehended ("explained"). He is concerned with the fact that logically simple explanations of the type "if A then B, and

[27] The theorems mentioned above concerning limitations of mathematical logics and the theory of computability do not contradict this point, since every *single* problem can be handled by outgödelings (as was shown above).

therefore B is caused by A" are insufficient for the analysis of complex system processes, and therefore one has to describe the system's causality (which is in fact presupposed) differently - that is, in a "negative" way. This is not least because of the logics of dynamic systems, which have been analyzed several times above, in terms of feedback, non-linearity, interdependencies between rules of different logical levels, causes that become their own effects, and Hofstadter's strange loops. In the face of all this, attempts to reach propositions of the simple type "from A follows B" seem rather hopeless.

Bateson could not be aware of the systematic modeling of complex systems with computer programs carried out to the extent that it has now become the heart of an established line of research. So his early insights are all the more admirable, since the computer modelings we are dealing with in fact proceed in exactly the same way: First, a model is constructed in which very much should be "possible" because of its logical extension. This means that one designs a space of possibilities for which the real processes form only a part - usually a very small one. In the next step the formal properties of the model are analyzed, so the question is here: On what conditions does the model behave in a specific way? Up to this point, everything happens, or can happen, without empirical "ingredients": Specific system parameters, rules, and meta-rules are varied, and the *fundamental* behavior of the system is analyzed. At this point, it is already possible to classify roughly the real system in a particular class of model behavior. This will be demonstrated in chapter 3. Finally (and this is the decisive step), constraints that are known empirically or from the theory of the real system are introduced into the model, which normally radically reduces the space of possibilities. Specific simulations result from this procedure, and thereby the comprehension of the model's dynamics, which determine the adequacy of the model. If the model is adequate, the constraints "explain" the observed processes, as Bateson notes, and determine what is not possible at this point. However, this is an explanation that only refers to the reduction of the space of possibilities and to the presupposition that the real system behavior is part of the constructed space of possibilities. Let us examine some examples to clarify this.

Together with Christina Stoica I have modeled the thesis of Eder and Habermas concerning the development of class societies by employing an interactive neural network (Klüver 1995). The model was designed in the way that the network consisted of 11 components - artificial neurons - with 11 * 11 = 121 weighted connections between them. If one allows the possible values of the weights to vary between (-1) and (+1), and allows only 10 possible values because of practical simplification, e.g. -1,

-0.7, -0.4, -0.1, 0 etc., then one gets a total of 10^{121} possibilities, and thus a figure with 121 zeros. So, in the first step a space of possibilities is given that is mathematically unimpressive, but that is nevertheless practically almost unmanageable, and that in particular contains all the possibilities that could never be realized (for whatever reasons).

The second step is not important here, because the formal qualities of interactive networks are quite well-known. The third step, however (that is, the introduction of constraints), is all the more important, because this is the only way to reduce the space of possibilities to such a degree that it is possible to check the model in a reasonable manner. As shown in Klüver 1995, this was done on the basis of historical facts and theoretical assumptions. We presupposed that a slight social differentiation had already taken place. As a result, the weights between the *social* units (the families) had to be asymmetrical, and had to be determined as decreasing or increasing sequences. In addition, we assumed on the basis of the processes of Christianization in Northern and Central Europe that higher social levels had a positive interest in the introduction of new religions and social roles, while the lower levels stuck to tradition. This again resulted in particular constraints regarding the weights - e.g. positive weights in both directions between the socially high families and the units of the new religion and the new roles. Finally, we determined that the units were not autocatalytic, but gained their increase only by interacting with other units. That was a plausible, albeit debatable assumption which in this case established that the weight matrix would have only zeros in its main diagonal. With these constraints it was relatively easy to construct a weight matrix that generated the required simulations.

A much more significant example can be found in Kauffman's considerations (1992 and 1993) concerning the role of self-organization in biological evolution. In the late sixties Kauffman, inspired by the studies on the differentiation of cells by Monod and Jacobs, had already begun modeling the cell as a kind of logical network. For this purpose he constructed an artificial net whose units represented the genes of a cell. However, he did not choose weights to represent the links between the artificial genes, as in artificial neural nets, but used instead so-called Boolean functions. These are simply the logical bonds of propositional calculus, namely the logical combining of propositions with "and", "or", "not", "if..., then...", and so on. Kauffman chose this way of modeling because he defined the links (and thus the rules of interaction) between the genes as the "switching on or off" of particular genes by particular other genes. This is the logical quintessence of Jacobs and Monod's Nobel prize winning discovery of so-called "regulatory genes."

First of all, this sort of model also constitutes a very large space of possibilities: Since every gene is either switched on or off at a particular state of the network (i.e. each gene has exactly two possible states), there are 2^n possible states of the net when there are n genes. Kauffman worked with networks of hundreds of "genes" - with 100 genes there are already about 10^{30} possible states. The number of possible Boolean functions, that is, the rules of interaction, is correspondingly extensive: If a gene is switched on or off by a total of k other genes, and thus k genes simultaneously determine the gene's behavior, then there are $_2 2^k$ different Boolean functions altogether. With k = 10 there already results a figure with about 300 zeros. (This is a classic example of the significantly greater number of possible rules of a system in comparison with its possible states, which was mentioned in connection with the concept of "memory".)

Kauffman knew from the studies by Jacobs and Monod that k could not be very large. Therefore, it seemed promising to take the values of k as the constraints. He then experimented with different values of k, and came to the following conclusions regarding the behavior of his network: with k = 1 (or less in the statistical mean) the net did not develop (or did so only slightly). It achieved certain states and retained them. With k = 2 different attractor states ensued for the network (depending on the initial states). The program ran through different cycles and stabilized at one of the attractors. With larger values of k the net became chaotic - that is, it did not show any discernible regularities and did not achieve any attractor states.

For Kauffman the question was whether (and how) the number of possible attractors was connected with the number of genes. This question was important to him because when he constructed his model he began by defining the different attractor states in the model as different types of cells: All cells in an organism have the same genome, and thus the difference in the eventual systems depends upon which cells are actually switched on or off, and especially upon which states are attractors for them. A fully developed cell differs from another one (e.g. a liver cell differs from a lung cell) in that it corresponds to a particular attractor state of its genetic network, which means that it does not change much anymore, and that the other cell corresponds to another attractor state. Thus, the number of a network's attractor states is identical to the number of different cell types that the genome can produce.

To his surprise, Kauffman found out by experimenting with his artificial network (2^{nd} step) that with k = 2 the number of attractor states corresponded approximately to the square root of the number of his "genes" - that is, for 100 "genes" there were about 10 "cell types".

Kauffman found exactly these values in empirical studies on "real" organisms, where these values could not be explained. He concluded that his model could possibly explain the empirical results by the fact that a developing network of genes must always take these values *for mathematical reasons*. The only alternatives are either chaos (that is, no ordered development of an organism at all), or stagnation (which also entails the absence of differentiating development).

These computer experiments are remarkable in two respects: first, their precise predictions obviously confirmed the validity of the model - that is, the assumption that a cell can be depicted as a *logical* network of genes. In this regard Kauffman followed the classical approach of McCulloch and Pitts (1943), who were the first to examine logical networks. Second, however, it was demonstrated that the system's constraints were not mere empirical facts, which could just as well have been different. The development of a dynamic network that is capable both of differentiating distinct cell types (attractor states) and of avoiding chaos is only possible within the magnitude of k, which has been empirically ascertained. Kauffman's experiments with his artificial net demonstrated that the empirical facts in this case are determined by *formal* characteristics of that type of dynamic systems. The facts are not contingent, but instead necessarily result from the logic of dynamic systems. It is no surprise that Kauffman speaks of a "Kantian stance" in this context: "What must organisms be such that they can evolve adaptively?" (1992, 179). Obviously, such studies of "pure" (namely formal) systems analyze the conditions of the possibility of complexity - in this case that of the evolution of ever more complex forms of living systems.[28]

This Kantian position has been expressed even more pointedly by Maynard Smith. By means of the mathematical theory of games he analyzed which proportion of the two sexes of descendants in a biological population is the most favorable. *Mathematically*, he came to the conclusion that with regard to evolution a ratio of 1 : 1 is most stable. This ratio can also be observed empirically: on average, an equal number of both sexes is produced with the division of X- and Y-chromosomes. Smith remarks on this point: "But if the evolutionary stable sex ratio was *not* 1:1, I have no doubt that some other mechanism of sex determination would have evolved" (1993, 12). Thus, the conditions of the possibility of evolutionarily stable heterosexual populations lie in the mathematical regularities of such systems. At the same time, this explains the empirical fact of the ratio of 1:1. We shall repeatedly return to this matter in the following chapters.

---

[28] Actually, things are not this simple, as will be shown in section 3.1. Simple trajectories arise with $k > 2$.

Let us return to Kant once more. The examination of formal systems (as will be systematically described in the following chapter) with regard to its consequences for the behavior of real systems, as in Kauffman's example, obviously betrays a deeply Kantian interest: The goal is not primarily the investigation of particular real systems, but the analysis of the conditions of the possibility of complex systems in general. This sort of inclination can be derived literally from Kant: "Die Bedingungen der Möglichkeit der Erfahrung sind auch die Bedingungen der Möglichkeit der Gegenstände der Erfahrung", is one of his most famous assertions (Kr. d. r. V. A 158. The conditions of the possibility of experience are also the conditions of the possibility of the objects of experience.). If we take this claim methodologically instead of epistemologically and apply it to the theory of models, then the conditions of the possibility of complexity (namely the formal properties of "pure" systems) are also the conditions of the possibility of the exploration of complexity: the *a priori* study of precisely these pure systems in the computer, as it were. Incidentally, this development also corresponds to Kant's well-known esteem for mathematics, that " pride of human reason" (*Stolz der menschlichen Vernunft*) (Kr.d.r.V. A 445). Only if knowledge can be formulated mathematically it is scientific knowledge.

As a conclusion to this chapter, let us turn to a problem that has been much more frequently discussed, namely the question of the "functionalism" of systems theory. This term usually carries two implications: First, there is the assumption (or presupposition) that systems can only be explained by interpreting their essential features as functional performances in adapting to their environment. As far as I tell, this thought was set forth systematically for the first time in classical Darwinism, and had begun to be employed to the benefit of the social sciences at least by the time of Malinowski's famous ethnological studies (1944). Second, functionalism is conceived as a kind of antithesis to causal explanation: If the latter recognizes a realized state of a system (in connection with the rules) as the cause of the following states and describes this relation as an explanation for the system's behavior, then a "functionalist" explanation, in contrast, would entail that the cause of the system behavior lies in the system's orientation to its environmental requirements, and thus to *states that have not yet been realized.* An example would be a society's introduction of universal conscription in order to cope with *future* conflicts with other societies. Of course, these two conceptions are closely connected, since instances of adapting to an environment are often tied to anticipations of the possibilities of one's own behavior. To simplify matters, however, I will treat these perspectives separately.

It should be evident that the states of an adaptive system at a particular time always have to be interpreted in part as the results of its interactions with its environment, that is, its evolution in regard to a specific environment. In this respect, the first conception of functionalism above simply points to the fact that one cannot understand the dynamics and the evolution of a system solely as isolated system processes. Luhmann's strong emphasis on self-organization (conveyed by his adoption of the concept of autopoiesis) sometimes leaves this aspect of system dynamics in the background. Nevertheless, adaptation is, of course, a decisive factor for the system dynamics of social systems too. We shall look at this claim in more detail in the coming chapters. It would be wrong, however, to be taken in by a functionalist version of the reductive fallacy (as has often been the case in so-called functionalist sociology). It is on no account appropriate to derive every specific property of a system, or every system state, respectively, directly from functional instances of adaptation. This inclination has to be ruled out for at least two reasons, to which I will confine myself here.

First, as I already pointed out in section 2.2., the self-organization of a system is, of course, always a formative factor for system dynamics. The environmental requirements can only tell *if* the system has any deficiencies with regard to its requirements. The question of *how* the system can rectify these deficiencies is completely up to the system's own logic, dynamics, namely its rules of interaction and the meta-rules that vary them. One might say that the system's environmental requirements posit an abstract space of possible developments in which the system has to realize its actual coping with these environmental requirements for itself. Since this usually leaves the system rather significant room to move, a real system state can almost never be explained directly from the system's environmental requirements, but only from the "history" of the system - that is, the interdependence between the rules on different levels and (most of all) the variation of the rules of interaction. In particular, the room to move makes it possible for the system to look not only for those own solutions which are more or less compatible with its previous history but also to hand down or introduce "dysfunctional residues" that have nothing to do with coping with the environmental requirements. This stems from the fact that systems practically never look for the "optimum" solution, which often does not exist for purely mathematical reasons, but content themselves with a fairly adequate solution. Such a solution can be mathematically sub-optimal, and in most cases actually is so, but is nevertheless sufficient in practice. Such an adequate though sub-optimal solution always makes it possible for a fairly complex system to retain dysfunctional residues. Another system may find a different adequate

solution for the same requirements, and then hand down other dysfunctional residues accordingly.

Therefore, "functionalist explanations" can only mean that the history of the system is interpreted as the self-organizing and adaptive variation of states and rules, the result of which may be neither unambiguous nor optimal. Thus, an appropriate analysis of the system must include both the reconstruction of the history of the system *with regard to the environmental requirements* (where it will especially have to keep an eye on the permanent interdependence between self-organization and adaptation), and take the uniqueness of the system's solutions, including their dysfunctional parts, into account.

A second feature of the history of the system is that it does not consist merely of the application of rules of interaction to initial states and the variation of those rules with respect to environmental requirements. The biological concept of mutation points to the fact that the history of the system also includes coincidences that cannot be explained by the interaction of different level rules. As a quantum physicist, Gell-Mann is, of course, quite familiar with this concept. He speaks of "frozen historical accidents" in this context (1992, 9). What he means is that the real history of a system is often influenced by events that are coincidental *for the system*, that yet somehow "freeze into" it. Despite their accidental nature they remain, and as a part of the system states they take part in determining its further history. Individual inventions might serve as examples for social systems: The invention of printing from movable type by Gutenberg may have been an event as equally coincidental as the prior invention of this technique in China and Korea. Only the system rules can explicate how that event is received by the system - Europe is known to have taken it differently than China.

It must be mentioned, however, that on a deeper level such "frozen accidents" might not be coincidental at all. Perhaps, for instance, a detailed biography of Gutenberg would show that he became an inventor and made his discovery with a certain necessity. On the social system's level, however, the invention is coincidental, and that's what matters here. As a consequence, we always have to take into consideration the fact that specific characteristics of a system might well be the result of chance. This fact, along with the system's freedom to realize its own specific and adequate, though suboptimal, solution in the space of possibilities of its environmental requirements, limits the applicability of any functionalist reductive fallacy.

The second meaning of functionalist explanation has to be dealt with in a more complex fashion, especially because of its obvious correlation to the differentiation of "because-explanations" and "in-order-to

explanations" that has been accepted since Schütz (1971). The "because-explanation" corresponds to a classical causal explanation in that the behavior of a person (a psychic system) is described as the effect of former events. On the other hand, an "in-order-to explanation" refers to the actor's intentions, and therefore to something that simply has not happened yet, namely something that is desired or averted - that is, something that may never become a reality. Accordingly, a functionalist explanation, which measures a system's behavior on its assessments of the environment and the strategies that result from that, would have to derive an "explanation" from the potential future of the system. However, a (natural) scientific causal explanation could not suffice anymore to explain this sort of process.

Since we are going to deal with this problem later in connection with the model of hybrid systems, I can confine myself here to a brief description of "anticipating" systems, where the same difficulty shows up quite clearly. Actually, anticipations of one's own behavior are simply analyses of the type "what would be if". In the course of such analysis, the possible strategies are evaluated with regard to the presently known environmental requirements, and (on the basis of this evaluation!) are selected with regard to the probability of their success. So in this respect, there is indeed an orientation to the future, though not the one that actually comes in reality, but to a "virtual" future, namely one that is constructed as a *model* by the system itself as part of its *present* state. Thus, strictly speaking, when the system anticipates, it only orients itself toward itself, namely to the parts of its state that it has produced in advance as anticipations, while it realizes the "actual" future in the process of its ongoing dynamics. In particular, the evaluation criteria that serve as the basis for the selections of an actually adopted strategy are themselves results of the history of the system: They are part of the experience that specifically learning systems have made with their environment, or - e.g. in the case of instincts - an essential part of the rule system.[29]

Of course, it makes a significant difference whether systems merely respond to environmental events (with or without changing the rules of interaction), or whether they instead behave "strategically," employing anticipation. Just because one cannot simply identify the one with the other, it seems quite reasonable (especially for the analysis of social action) to maintain Schütz' differentiation of "because" and "in order to". It simply must be perfectly clear that one is not taken in by the

[29] This becomes clear with the example of hybrid programs in chapters 3 and 4:entire programs construct possible futures that are at the same time part of their present "real" state. The actual dynamics of the formal program systems ensue afterwards from the decision for one of the possible futures as a substate of the system.

misconception that "in-order-to explanations" are incompatible with strict systems theory, and, in particular, that they are not a form of causal explanation. This would constitute a reductive fallacy for the humanities in the sense that it would reduce complicated cases of systems theoretical causal analysis to a simple form of commonplace hermeneutics. In fact, "in-order-to explanations" are causal explanations too, in the sense that the behavior of a system is derived from the states that the system has already realized. However, they are explanations that cannot be expressed in terms of simple if - then relations, but only in the somewhat more complicated logics and mathematics of layered rule systems. In particular, it is no problem to reconstruct these explanations formally.

Regarding content, this argumentation is, of course, not new. Rather, it recapitulates in terms of systems theory an argument that is as old as the attempt to put the "cultural and cognitive sciences" on par with the precise natural sciences. My point is not to try to tell e.g. phenomenologically orientated social scientists not to be concerned with Schütz' analysis anymore - on the contrary. I have instead attempted to show that it is perfectly possible to integrate the different aspects of human action consistently into the formal logics of the terminology and methodical procedure of systems theory, which means that the complexity of social action and the behavior of social systems do not comprise limits for the scope of the scientific approach that has been presented here - no more, no less.

3. THE DYNAMICS AND EVOLUTION OF FORMAL SYSTEMS

> "The dynamical systems we are going to investigate are not models of particular physical, chemical, or biological systems. We are interested in the generic properties of the dynamics of self-programmable matter. By implementing the property of self-programming in a simple computational system we hope to be able to extract such generic properties. Hopefully the dynamics of the emergence of novel properties in such a simple system also enables us to understand aspects of the emergence of novel properties in the more complex natural systems." (Rasmussen et al. 1992, 214)

In the first two chapters I repeatedly pointed out that the construction and analysis of computer programs currently serve as objects of systems theory research, and explained why this is so. Here I would like to sum up systematically once again the several reasons for this state of affairs.

On the one hand, computer programs simply are the most important (cognitive) tools for the sciences described here though certainly only ones. Therefore, a precise knowledge of these tools should be just a matter of course, particularly because they are not in the least trivial. Of course, one can leave the technical problems of the computer as a material tool to competent specialists, as I do all the time. However, one should know the programs' properties as *theoretical* tools, much as a theoretical physicist has to understand the mathematics of differential equations.

On the other hand, however, computer programs of the kind we will deal with here are more than "just" tools. They might be described as "pure" systems by means of which the behavior of complex systems can be analyzed "per se". In this context Stanislaw Ulam, one of the originators of cellular automata, has said they realized an "imaginary physics" - the study of dynamic physical systems in its pure form (quoted after Vichniac 1984). Thus, the construction and analysis of computer programs is related both to the experimental methods of the natural sciences and to their mathematical theory formations: as in an experiment, an idealized reality can be examined and manipulated without disturbing or uncontrollable influencing factors; as in classical mathematical theory, individual results can be combined in a formal model and can be "explained" by this means. This is so particularly because the programs in question are always designed as models of complex dynamic systems. Since they are "pure" systems they represent the *generic* properties of the systems for which they are a model and are themselves their own idealizations. Burks makes this perfectly clear

J. Klüver, *The Dynamics and Evolution of Social Systems*
© Springer Science+Business Media Dordrecht 2000

when he remarks with regard to John von Neumann's research interest for self-reproducing automata: "Von Neumann [...] was not trying to simulate the self-reproduction of a natural system at the levels of genetics and biochemistry. He wished to abstract from the natural self-reproduction problem its logical form." (Burks 1966, quoted after Koza 1994, 226).

Of course, this does not mean that the analysis of such pure systems already suffices, for example, for research on social systems. For instance, precisely because of its transparency, the fox and goose system that was sketched in chapter 1, was idealized so radically compared with real predator-prey systems that we cannot *directly* infer the behavior of real ecosystems from the behavior of the program. It means instead that now, with these programs, a concrete meaning can be assigned to the concept of a general systems theory: one can design and analyze systems that are indeed general, because they demonstrate the possibilities of system behavior in general. In other words: just as the limits of the possibilities of mathematical natural scientific theory formation are the limits of the formulations they use, the limits of the behavior of pure systems demonstrate the limits of what can possibly be expected in reality. At the same time, however, the possibilities of pure systems demonstrate the possibilities of real systems - even if these possibilities have never (yet) been observed in reality.

This is a consistent continuation of the "Kantian stance" mentioned in the previous subchapter. Precisely as Kant is known to have seen mathematics (i.e. the sphere of synthetic judgements *a priori*) as the science of pure intuition, and thus of the pure conditions of the possibility of scientific knowledge, the analysis of pure systems can be understood as the condition of the possibility of the exploration of real systems. Just as for Kant, empirical observations remain indispensable because there is nothing that can be inferred *directly* from the analysis of pure systems for the behavior of physical or social systems. We can, however, learn from such analyses what can be known in principle about real systems, and what we must pay attention to when we analyze real systems. At the same time, the mathematical regularities obtained from the analysis of pure systems explain the facts (or the empirical knowledge) that can be observed in the real systems for which the pure systems are models. We could note this fact quite clearly in section 2.7 in the results of Kauffman and Maynard Smith.

If we formulate this view in a more sociological fashion, the analysis of pure systems obviously offers the following possibilities: At the beginning of this study I pointed out that a sociological systems theory can only be justified if social systems are (necessarily) conceived as the consequences of rule-governed actions or interactions. The focus of

theoretical interest, especially for sociologists, thus shifts from the "original" system dynamics to the corresponding rules: What properties must rule systems possess in order to make the behavior of complex systems explainable? In other words, the issue is not primarily *how* specific systems (or system classes) behave. This is, instead, a derivative question. Our main concern is with the properties of rule systems from which the system behavior results necessarily. Thus, the fundamental question with regard to the analysis of "pure" rule systems is whether there are universal characteristics of rule systems to which every single concrete rule system is subject. In the following sections this question is dealt with through the search for so-called control parameters or meta-parameters respectively. However provisional most of the results presented here may be, some basic trends are already becoming apparent. In spite of all their heterogeneity, rules and meta-rules are not arbitrary, but follow general patterns and trends, comparable to those of heterogeneous, distinct languages. With reference to Chomsky, it is certainly not exaggerated to speak of possible *rule system universals*.

Generally speaking, pure systems play a technical, a transcendental, and a heuristic role for each systems theory. Inasmuch as they also explain which observable system behavior can be understood by these theories, they even function as explanations. Let us therefore take a closer look at them.

3.1. CELLULAR AUTOMATA AND BOOLEAN NETS: THE PARADIGM OF SELF-ORGANIZATION

3.1.1. The Logic of the Chessboard: Cellular Automata (CAs)

> "The beauty of the CA approach is clearly its simplicity" (Rasmussen et al. 1992, 219).

For the description of dynamic formal systems I will concentrate in this section on CAs and a relatively short parallel description of the so-called Boolean nets. There are several reasons for this concentration, which also underlie both the fact that we systematically experiment primarily with CAs at the University of Essen (in combination with genetic algorithms), and that simulations with CAs also are increasingly applied in the social sciences (e.g. Hegselmann 1996; Nowak/Lewenstein 1996; Klüver 1995; Passerini/ Bahr 1997). This focus on CAs is not a limitation to generality, however, since it can be shown that CAs with appropriate rules are

"universal" automata. Logically, they are equivalent to a universal Turing machine, which means that any formal systems can be simulated with an appropriate CA. In other words, this means that any computable function can be computed with a CA - although such modeling might be somewhat involved. For the purposes of simulation this means that any (physically) real dynamic system that can be described formally can be simulated with a CA.[30] These are the reasons why CAs have been thoroughly analyzed, particularly as dynamic systems *sui generis* (Wolfram 1986; Langton 1992; Rasmussen et al. 1992; Wuensche/Lesser 1992). Hence it makes sense to take CAs as examples of pure systems that are especially well examined and understood.

The primary advantage of formal modelings on the basis of a CA is that they particularly facilitate analysis of the system dynamics generated by rule-governed interactions on the basis of the bottom up approach described in section 1.2. The CA's logical starting point is always the construction of local rules of interaction (which is also true for Boolean nets). The central question, then, concerns which rules of interaction produce which dynamics. Therefore, the sociological research interest in pure systems that was sketched above can be realized extremely well through the use of CAs. If one begins with the rule concept as the central social category (for which in the case of social systems there is often no other way for strictly technical reasons - see below, chapter 4 and section 5.1.), then traditional modeling techniques for describing system behavior on the basis of equations are not suitable. By comparison, models such as the CA, which operate on the basis of rules, are in a sense the natural formal counterparts of social systems, (which are also based on rules).

Vichniac expresses the basic properties of CAs as follows:

> "In fact, cellular automata have enough expressive power to represent phenomena of arbitrary complexity [...] To appreciate the originality of this type of modeling, one should keep in mind that there is no attempt here to solve any given equation: in fact cellular automata [...] merely perform simple space dependent logical decisions." (1984, 97)

This expresses exactly - from a physicist's point of view - what the sociological interest in "pure" systems is all about: it cannot be the

---

[30] To put it more precisely: the so-called Church-Turing hypothesis says that no formal system is more general than a universal Turing machine, or its logical equivalent respectively. This hypothesis has not been proven yet, but practically no logician or computer theorist really doubts it. Accordingly, the *physical* Church-Turing-hypothesis says that any physical process - any physical system - can be precisely described with a universal Turing machine. For non-computer scientists it might be useful to add that universal CAs can also be seen as universal "virtual" computers, that is, as simulations of physically real computers in the computer. In Rasmussen et al. (1992), there is a short list of universal formal systems.

(usually fruitless) search for general equations which describe the system behavior "globally". It instead concerns the analysis of rules and the consequences of following them - "simple logical decisions". Such a procedure is as unusual for physicists as it is (or at least should be) natural for sociologists.

There is yet another advantage: most traditional mathematical techniques, such as partial differential equations, probability theory or Markov processes assume more or less explicitly linear and/or additive phenomena. Dealing with non-linearity usually requires linear approximations. "CAs constitute a notable and interesting exception" (Holland 1998, 232) because they allow the analysis of non-linearity in a "pure" form. Since social systems are primarily non-linear systems, CAs are especially well suited for the analysis of this property of social reality. Together with Jörn Schmidt, I have already described some of the possible sociological uses of CAs in (Klüver 1995). In addition, the following description suggests a stricter systematic for CAs and in particular delves into the analysis of so-called "Wolfram classes", as well as the most important control parameters:

A CA program produces a grid of "cells" - either a fixed or an arbitrary number, depending on how long the program runs. The screen is divided "phenomenologically" into a grid of cells, which are usually marked by different colors and/or geometric figures. Every cell can take a finite number of states. The simplest case (which is realized, for example, in the famous "Game of Life") consists of two states: on or off, 1 or 0, alive or dead. This is why such cells are often described as "finite state machines". The state values are usually represented as integers, that is, whole, positive numbers. Of course, it is also possible to represent states as multi-dimensional vectors. We have often used this method in sociological simulations. Thus, what we are dealing with are systems whose elements can take different states themselves, in a way similar to the already mentioned neural nets.[31]

The rules according to which the cells' states or state values are determined are usually "local" rules - that is, the state of a cell results from the states of its neighboring cells. In two-dimensional CAs (which are used mostly for simulations) the cells are displayed as geometric squares. The "neighboring cells" are either the four cells at the sides of the square or, those cells in addition to the four cells at the corners (i.e. eight cells in total). The neighboring cells are described as the "neighborhood" of the computed cell. In the case of four cells we have a von Neumann neighborhood (after the inventor of the CA), whereas in

[31] For the sake of simplicity, one can of course also see every state of an element as a distinct element itself.

the case of eight cells we have a Moore neighborhood. One can, of course, also create neighborhoods of six, three, or five cells, but this would not in principle increase the possibilities.

One-dimensional CAs have also been investigated, especially by Wolfram (1986). These are, however, of interest primarily for methodological reasons. One can also use three-dimensional CAs, as has already been done in the Game of Life (Dewdney 1988 a). Here the combination possibilities are, of course, highly increased, since, for example, a Moore neighborhood that consists of eight cells in a two-dimensional space has 26 cells in a three-dimensional space. In general, Moore neighborhoods consist of $3^n - 1$ cells in an n-dimensional space.[32]

This extreme simplicity of the basic structure of CAs allows almost any kind of variation, which provides the fascination of these formal systems. In particular, it is quite feasible to investigate the behavior of formal dynamic systems by means of systematic rule variation. The very fact that in the case of k state values and n neighboring cells a program has k^n rules and ${}_k k^n$ possible sets of rules demonstrates the vast range of possible rules. As one can see, one can soon encounter "combinatory explosions", which produce the great variety of possibilities for CAs.[33]

Usually, one does not bother listing every single rule (that is, every possible neighborhood combination) when constructing a CA. For

---

[32] To elucidate this point, consider the following: obviously, you get from n = 1 with two neighboring cells to n = 2 with eight neighboring cells by adding two more cells at the top and at the bottom of each of the two neighboring cells and in the gaps, that is, at the top and at the bottom of each corner. You can carry out a corresponding procedure to get from a two-dimensional neighborhood with eight cells to n = 3 with 26 cells. This results in the recursive *sequence* $x_n = 3 * x_{n-1} + 2$. As with any proof by induction, we can assume that the above dimension formula is already proved for n-1; as a consequence, we get for any n $3*(x_{n-1}) + 2 = 3*(3^{n-1} - 1) + 2 = 3^n - 3 + 2 = 3^n - 1$. (The idea for this form of expression was inspired by Mirella Skakic and Jörn Schmidt.)

[33] For those with a mathematical interest: You can easily understand the above formulas for k = 2 if you see them as special cases of a more general theorem of set theory: a set of n elements always has a total of 2^n subsets (the so-called empty set, which consists of no element, and the set itself are counted in). When the number of rules for a one-dimensional automaton is determined, a triple (a,b,c) is interpreted as an empty set, if all values are zero; if all values are 1, then this triple itself is the set; the other combinations are handled accordingly. Evidently, this results in a number of 8 combinations. The sets of rules that consist of 8 rules are treated in the following way: the set of rules of the computed cell that takes the value 0 for all rules is defined as an empty set, and the set of rules that takes the value 1 for all rules is defined as a set itself. A set of rules that takes the value 1 for one rule and 0 in all other cases is a subset with only one element, and so forth. The result of this is the above formula. The general formula is then achieved by induction over k.

example, in the case of social system simulations it is usually sufficient to work with mean values of neighboring cells, which significantly reduces the number of necessary rules. In these cases we speak of "totalistic rules". On the other hand, one often needs more than two state values, and the neighborhood does not necessarily consist of the immediate neighboring cells. These facts, of course, considerably increase the number of rules and the possibilities of their combination. This will be clarified later by examples.

The logical structure of a CA can also be described more formally, using "k" as the number of cell state values (the basic components) and "n" as the neighborhood quantity (the "degree of computation"). Since the computations of the CA can also be interpreted so that individual rules indicate the type of single cell transitions from one state to another, and since this happens on the basis of the rules that are defined by the combinations of the neighboring cells' state values, one can conceive a specific set of rules (a program) as a transition or system function f. Thus, the transition function f maps a number space (whose dimensionality n equals the size of the neighborhood plus the cell that is being transformed) into a single number - a scalar (if the cell's state value is only a number). In classical algebraic terms, the transition function is a mapping from an n-dimensional space into the one-dimensional space of the state values of the single cells.

In the case of k state values and a neighborhood size of n, there are k^n possible rules for the program. If we also assume k as the dimension of the system's state space, then the transition function f consists of k^n individual "transitions" and is a mapping of the k-dimensional state space of k^n combinations of single cells - the neighboring cells - into the state of a single cell. From a mathematical perspective, the running of a program are obviously merely the repeated iteration of the transition function.

Basically, these logical constructs are rather uncomplicated formalisms. Accordingly, they can be written out as programs in a formal language. However, they are in fact more than the traditional symbolisms of logics and mathematics, since, as operational instructions for physically real computers, they develop in space and time (which is to say, they are perceptible in space and time). I have already indicated this point in section 2.5. It is particularly conspicuous in the case of a CA, since the succession and variation of its different configurations is immediately observable as a development of geometric figures. So in a way, CAs are spatio-temporal (i.e. dynamic) realizations of the pure mathematical and logical structures. In particular, they often develop a system behavior that could not have been foreseen by intuition when the

program was constructed - and is often not objectively foreseeable. This is the case whenever the rule system of a CA represents an analytically insoluble equation system. Hence, it is more than just a metaphor when I employ the concept of dynamic systems *sui generis* (i.e., "pure" dynamic systems) here.

This becomes especially clear when we look at the results of Wolfram's (1986) experiments with one-dimensional CAs. Together with Packard, Wolfram was able to generalize these results for two-dimensional CAs.[34] Wolfram experimented with numerous CAs. He restricted the number of state values to a maximum of five and assigned the size of the neighborhoods accordingly. He discovered through systematic rule variation that CAs may be divided into four classes according to their behavior, the so-called Wolfram classes:

Class I develops (quite quickly in most cases) towards a fixed final state. If that state is reached, it remains constant. There is a sociologically inspired example of this in (Klüver 1995): The development of class struggles according to Marx is simulated in such a way that after an average of 200 runs the program reaches its final state of (generally) achieved socialist revolutions. Thereafter there remains a society of proletarians with a high class consciousness. Hence, we are dealing here with systems whose dynamics can be characterized by a point attractor (period 0).

Class II develops relatively simple geometric patterns, which do not, however, remain constant. They disappear in favor of other geometric patterns and reappear after certain periods of time (which are usually quite short). On the whole, we can characterize this developmental process by saying that the systems of class II have several periodic attractors. After a short time, one of these attractors is reached (depending on the initial state), and thereafter the system only runs along the path determined by the attractor. Given specific parameter values, the repeatedly mentioned fox and goose system can be described as a class II system. When its spiral patterns occur, the system merely "oscillates" around the attractor.

Class III consists in the generation of completely aperiodic patterns whose sequence can not be ascribed to any regularity. This behavior is obviously chaotic. In fact, the system is in a deterministic chaos - all the rules of this class (as well as of the other CAs Wolfram has experimented with) are deterministic. Attractors, if they exist at all, only exist as

---

[34] I have not read about any corresponding results regarding three-dimensional CAs yet; however, we assume from our present experience with three-dimensional CAs that the results would be formally the same. This assumption follows from the so-called control parameters of the CAs' rules systems, which are described below.

strange attractors. Examples with graphic illustrations can be found in (Wolfram loc. cit.) and (Langton 1992).

Finally, class IV does not behave chaotically. It is, however, much more indistinct than the classes I and II. The systems of this class generate complex structures that recur in most cases only after "very long transients" (Langton 1992, 46). The behavior of the systems of class IV - although it is ordered "in principle" - is practically impossible to predict. The most famous example of a class IV CA is probably the "Game of Life" by Conway (cf. Poundstone 1987; Berlekamp et al. 1982). Packard and Wolfram assume that special problems of computability (so-called "NP-complete problems") and, in borderline cases, problems of undecidability (see above, section 2.5) particularly occur in this class, so that the behavior of these systems can actually only be recognized by actually running the program. In particular, universal automata, for which the Game of Life is the first example, only occur in this class, according to our present knowledge (Berlekamp et al., loc. cit.).[35]

These four classes can therefore be described in the established terminology of the theory of dynamic systems as follows: The systems of class I develop towards limit points (and then stay there), which means that they only have trajectories with point attractors (period 0). Systems of class II tend towards limit cycles or simple attractors, and move periodically between the attractor states. Class III is characterized by the concept of deterministic chaos; the development of the system states is determined by strange attractors. In principle, these classes are already known from the physics of dynamic (continuous) systems. In contrast, it is interesting that according to Wolfram class IV has no direct analogy to continuous dynamic systems (CAs have discrete sequences of states). This is obviously something new for the general theory of dynamic systems that has developed from physics.

If we observe these classes according to their complexity, it quickly becomes apparent that they demonstrate a characteristic sequence of complexity values. Following a suggestion by Gell-Mann (1994), we confine ourselves to the combination of effective complexity (K_e) and algorithmic complexity (K_a) as a degree here (see above):

---

[35] This claim has to be somewhat restricted: *finite deterministic* systems are always periodic, i.e., they always have simple attractors. If a finite deterministic system starts from an initial state A, and if it has, e.g., n possible states, then it must re-reach one of the previously realized states sometime and run through the same closed trajectory from then. In the extreme, the system reaches n-1 different states from A and then comes back to A. Hence, chaotic behavior can only occur if the system is stochastic or potentially infinite. However, in most cases it is practically impossible to decide in the given time to which simple attractor the system's trajectory runs so that *for an observer* the system seems to be chaotic.

. Class I can be characterized formally by a relatively small number of states whose sequence ends in the attractor of the period 0. Thus, K_e just consists in the provision of the states including the attractor and the note that the sequence of states ends there. This can evidently be characterized by a relatively short description. Accordingly, K_a is relatively small, since such a system can always be generated by a simple algorithm. As a result, we have $K_I = K_{Ie} * K_{Ia}$, which is a small value.

Accordingly, a CA of class II consists in a rather short sequence of states, which reaches a certain attractor (depending on the CA's initial state) and enters the period of the attractor. This period is again a sequence of states, which begins anew when the attractor is reached and is then repeated as long as the automaton is observed. K_{IIa} and K_{IIe} as well as K_{II} are correspondingly small, though greater than in class I. altogether, we have $K_{II} > K_I$.

Class IV can be determined analogously, and it is easy to see that K_{IV} is greater than the complexity of the two other classes. Since these periods are rather long, the scheme of description is necessarily longer than in the first two cases so that $K_{IVe} > K_{IIe} > K_{Ie}$. A rule system that produces a class IV system is generally longer (or is at least not shorter) than those of the classes I or II. As a result, the systems of class IV are the most complex according to the computational formula for K.[36]

Things are different in the case of class III systems. Here we have a very high algorithmic complexity, since the sequences of states are for the most part arbitrary. On the other hand, $K_e = 0$ because there is no scheme at all for the description of regularities. Therefore, $K_{III} = 0$, and thus the systems of class III are the least complex. As a result, we get the sequence III - I - II - IV as the sequence of the increase of complexity. This corresponds to an intuitive understanding of complexity, at least inasmuch as complexity is seen as *both* order *and* "complicatedness" (that is, many-sidedness), etc.

By the way, if we had defined K as a vector length there would be a different sequence, namely I - II - IV - III, or I - II - III - IV respectively, depending on the size of K_{IVe}. This runs counter to intuition only if complexity is not associated with order (which is seldom the case). Therefore, it is completely justifiable if we define K in such a mathematical way here.

---

[36] To illustrate this: Simple rules of a rigid class oriented behavior of capitalists and proletarians were enough in the mentioned simulations to get a CA of class I. A "social democrat" variant resulted in a class II CA, in fact by adding extra rules that made the capitalists "able to learn". K_a was increased and with that K_e grew too. We got CAs of class IV by a further extension of the rules, for example by introducing charismatic leaders of the working classes and soldiers.

As we have seen, CAs demonstrate all the characteristics (and more) that are known as the fundamental behavior of dynamic systems. This is a strong, experimentally obtained confirmation of the noted universality of CAs, and a clear indication of the fact that one can really learn the basic general aspects of the behavior of dynamic systems from the analysis of CAs. Furthermore, there are strict proofs for the universality of single CAs, especially for the "Game of Life" (Berlekamp et al., loc. cit.). If we think of the CA's simplicity in principle, these results may seem quite surprising at first and certainly contradict intuition. What is the reason for this hardly comprehensible variety? In other words (and with regard to the construction of formal systems), how can systems of the separate classes be created systematically?

Some answers to that question have developed, and they all agree that single rule systems can be characterized by specific "control parameters" (Kauffman 1993). Since Langton's studies on the CA (1992) have become particularly well-known, I will pursue his analysis in more detail here. Wolfram's division into classes is merely "phenomenological" and only describes the surface structure of our formal systems, as it were. Accordingly, Wolfram has contented himself with statistical characterizations of his classes. An analysis of the "deep structure" (Langton) at work here immediately reveals the variety of the system behavior:

Langton "parameterizes" the space of the transition function by introducing the following λ parameter: consider again k states of the cells and n neighboring cells, as well as a specific transition function f, that is, a rule combination of k^n rules. Now we choose an arbitrary state s that is described as a "quiescent state".[37] There are r rules (= transitions) in the transition function f, which produce the state s for the actually computed cell. Additionally, there are $k^n - r$ transitions to the remaining k-1 states. Then we have

$$\lambda = (k^n - r) / k^n.$$

Let us look at this parameter with regard to some borderline cases of r. If $r = k^n$ then $\lambda = 0$; all transitions lead to the quiescent state s. If $r = 0$ then $\lambda = 1$. There is no transition to the quiescent state s. If all states are reached with the same frequency the number of rules for the actual states is r, and thus $r = k^{n-1}$.[38] Then $\lambda = 1 - 1/k$. Obviously, the parameter values $\lambda = 0$ and $\lambda = 1 - 1/k$ represent the transition functions that are the most

---

[37] Langton is not quite accurate in this part of his description: our own experiments showed that s always has to be the state that is most frequently reached by the rules.

[38] Hence, with k states the total number of rules is again $k * k^{n-1} = k^n$.

homogeneous and heterogeneous respectively. Therefore, it is obvious to concentrate on CAs with λ values between 0 and 1 -1/k.

Langton demonstrates the effects of the variations of λ for CAs with λ values between 0 and 0.75. The case of $\lambda = 0$ obviously realizes the simplest case of a class I CA, because the system immediately changes from arbitrary initial states (system states!) into the quiescent state s for all cells and thus produces a final system state that exclusively consists of cells in the state s. In general, we can say that low λ values produce systems of class I. If we gradually increase the value of λ, we get systems of the classes I - II - IV - III in succession.

This sequence seems to be significant. Langton - and others with him (Kauffman 1992 and 1993) - drew the conclusion that real systems do not show a system behavior of class IV until the "processes of self-organization" of life develop on the basis of purely physical-chemical systems. As a consequence, life could be formally defined as a system behavior of class IV, located between simple periodic and chaotic systems. This has to be qualified though, because even the complex systems of Wolfram class IV are not adaptive. If life can be formally characterized by class IV rules, then this is "only" the model of a single organism and not the modeling of evolution (see below, section 3.2.).

Admittedly, the scope of λ values that makes a system behavior of class IV possible is rather small in relation to the other scopes of values; this was somewhat concealed in the summarizing description above. Life (and other really complex systems) permanently balances in a small sector between overly simple periodic dynamics and the equally life-threatening confusion of chaos. Therefore, Langton gives his studies (1992) the vivid title "Life at the Edge of Chaos", a concept that is also used by numerous other authors in this context (Packard 1988; Crutchfield/Young 1990; Kauffman 1993).

If we look at the λ parameter with greater consideration for its "content", it is obvious that the decisive variable is r - that is, the degree of how often a certain cell state is realized by the transition function. The rather summarizing interpretation of this fact is that the dynamics of the system are the more varied the more states with the same frequency can be realized by the transition function. If all states are realized with practically the same frequency, chaotic behavior ensues. Life (and other complex systems too) would then depend upon keeping their own "inner" flexibility high, but only up to a certain limit beyond which there would be chaotic states.

At this point we can easily clarify the λ parameter from a sociological point of view. If we think of two different societies, one of which has strictly regulated and allocated access rules to specific social roles (as in estate or even caste structures), and the other of which in principle allows

each individual to change from any role to any other (as in modern societies), then the first society evidently has a lower λ value, and the second has a comparatively high one. From this one might expect that the dynamics of the first society is rather "static" – in other words, that it constantly moves within the scope of a simple attractor with a short (or no) period length. On the other hand, the second society tends toward Wolfram class IV, with a correspondingly complex dynamics (i.e. long attractor periods). It is important to keep in mind that this is only true in case of constant rules. The comparisons between modern and comparatively traditional societies should easily confirm these intuitive considerations. The often disturbingly great variety of the dynamics of modern societies results not only but also from the fact that they permit a high degree of social mobility - that is, they are based on rules with a high λ value.

By introducing the so-called "Z parameter" Wuensche and Lesser (1992, 39) have taken a different view on the CA. We will make a brief excursion into the "temporality" of the CA to understand what this is about:

The laws of classical physics and chemistry are known to be "t-invariant", i.e., they are reversible in time. Therefore, it is possible in principle in the case of deterministic systems to determine both their previous and their future states (I abstract here from the problems of chaotic systems and of thermodynamics; for the latter see e.g. Coveney/Highfield 1990). Interestingly enough, this is not true for most of the CAs: Their rules are temporally "one-sided" - that is, they do permit an injective computation of future states from an initial state in the case of deterministic rules, but allow no *injective* computation of previous states from a given one. This is why physicists have often doubted the physical applicability of the CA. However, although most CAs are irreversible, it is easily possible to develop reversible ones as well, as was shown by Fredkin and Margolus (cf. Gerhard/Schuster 1995, 110). This goes to show (especially for social scientists) that even these very simple looking formal systems do have something like "historicity", which makes them particularly relevant for the modeling of social systems.

As a result, irreversibility means that it is impossible to determine a previous state from the knowledge of the present (and the knowledge of the rules, of course) *unambiguously*. This allows us to define the Z parameter: It measures the probability of the cell states being settled, that is to say, being injectively determined at the time t-1, if these are known for the time t (i.e. the temporally nearest point); thus, Z measures the probability of being able to infer the state z(t-1) from the state z(t). As a

consequence, we can classify rule systems with regard to the question of how known, unambiguous states can be "computed back" to previous ones. Since Z measures probabilities, it varies between 0 and 1.

Let us look at some examples: if $Z=1$, one and only one state z(t-1) is possible, this results in Z(t) if computed "forward"; with regard to the whole automaton, this means that there is only one trajectory that can lead to the known state Z(t). This applies, e.g., to reversible automata. If $Z=0$, practically any value of the cell at the time t-1 is possible, and there are very many trajectories that lead to the known state - depending on the number of generally possible states of the CA. Practically, this means that our state is a point in the state space where all, or at least almost all trajectories meet; from there on, only one trajectory remains. If $Z=0.5$, many (but not all) trajectories meet in this state (the relation of the variation of Z and the variation of the intersecting trajectories is not linear).

These examples already show that Z can also include the Wolfram classes. There are indeed correlations similar to those pertaining to the λ parameter: if Z is between 0 and 0.6, the system stays in class I or II; this is because the different trajectories are quickly "standardized" and enter the same attractor, if there are any. Thus the system can take only few attractor states, and it reaches them relatively quickly. An increase of Z results in class IV and then class III on the basis of the same logic, because now different states z(t-1) do not have the same attractor in most cases, and so the dynamics take long, and sometimes chaotic attractors. The correlation between Z and λ (described below) makes these largely equivalent results easily understandable.[39]

So far, we have only talked about deterministic CAs, and in fact these systems are presently the only ones that have been investigated in detail; by the way, this also goes for the Boolean networks described in section 3.1.2.[40] However, in most cases it is neither plausible nor practically feasible to work with purely deterministic systems when problems of the social sciences are concerned. One reason is that one often does not know enough about the systems in question and their rules to achieve more than just assumptions on the probabilities. To compound matters, individuals just do not endlessly act in the same situations according to the same rules. The proposition that specific rules are valid in specific

---

[39] Wuensche and Lesser's motive for introducing Z is not the analysis of the temporality of CAs, but the "geometry" of the state space with the help of the concepts of the 'basin of attraction' and the 'field of basins of attraction' (see below). This is also an important aspect; however, I think that the temporality of CAs is even more important.

[40] Bar-Yam 1997 is one of the few exceptions; however, he deals with stochastic CAs only briefly from a thermodynamic point of view.

situations can never be more than just a proposition concerning probability, because one can never definitely exclude the possibility that the rules are broken (for whatever reason). Individuals are just not completely determined by social situations and the rules valid in them. Therefore, social systems - apart from very simple cases - can actually only be modeled as stochastic systems. Of course it is important to gain insights into deterministic CAs, but this is equally true for stochastic CAs. However, the following considerations upon this matter are still rather preliminary.

If we want to transfer the concept of the control parameters to stochastic CAs, we have to define the concept of the "stochastic rule" in more detail first; in the 2nd chapter this concept was only summarized with reference to probability values. Now this position has to be stated more precisely.

In connection with the CA, "stochastic" basically means that the transitions of the single cell states "branch" according to specific probability values. What this means is the following: CA rules of the kind we analyzed so far always describe the transition of one cell state into one and only one other cell state. In the case of deterministic rules (which as a whole represent an injective transition function) there is no other possibility. Stochastic rules, however, can also have the following form: given the neighborhood conditions U and the cell in state Z, there is probability p that the cell will transform into state Y; probability q that it will transform into state X; and probability r that it will transform into state W. In other words: This sort of rule defines alternative possibilities of transitions that are realized by the system according to probability values.

It is obvious that social systems can often only be modeled in this way. Individuals are usually not tied down to a single action possibility, but instead have options (see below, chapter 4., about the theory of games) which they follow according to differing probabilities. A prime example is the socialization process: its " regularities" (e.g. with regard to the specific sex or class) only imply that though an individual has various possibilities for fashioning its life, the probability of his or her realization of these possibilities differ considerably depending on his or her social background or sex.

In this context, it is fruitful to transfer our consideration to the λ parameter, or to the equivalent P parameter that is described in section 3.1.2:

I showed that the dynamics of a deterministic CA depend on the control parameters of the respective rule systems. If we add probability parameters to these rules, we have a stochastic system that is determined

on the one hand by its fundamental deterministic rules, and on the other by its probabilities. The deterministic rules just describe the transitions of different cell states into others (or the same). We can describe these rules by a "frequency matrix" $F = (f_{ij})$, with the element f_{ij} representing the number of transitions from cell state i to cell state j; if we standardize these integers as proportions in relation to the total number of all transitions that are possible by the rules, then F consists of elements with values between 0 and 1.[41] Accordingly, the probability parameters for the respective rules can be described by a matrix $P=(p_{ij})$, with p_{ij} representing the standardized probabilities of the transition $i \rightarrow j$, of course.

Now we can form an FP matrix = $(f_{ij}*p_{ij})=(fp_{ij})$, whose elements are the products of the single matrices' elements (which is not the usual matrix multiplication!), and which are a combination of the frequency *and* the probability of the different transitions. A control parameter that can characterize these "dualistic" rules - they consist of both deterministic rules and stochastic ones - merely measures the difference of the value distributions in the FP matrix, analogous to the λ and P parameter. If we call this parameter FP too, and if k is the state with the greatest value in the FP matrix by adding up all transitions into k, we get

$$FP = (\Sigma_i (\Sigma_k fp_{ik} - \Sigma_j fp_{ij}))/ (n-1)$$

with n elements in the FP matrix. Thus, what are being measured are the differences of the frequencies and probabilities of the transitions in relation to the greatest value; FP = 1 is of course just the deterministic borderline case with a simple dynamics.

Experiments of Ralph Kier have shown that FP is only a very rough measure; however, its use allows the discernment of some trends. Values of FP above 0.3 always result in a simple global dynamics, which is of course locally distorted; that is exactly what was to be expected from stochastic systems. Therefore, in analogy to deterministic CAs we may speak of "simple" and "complex" dynamics. More complex global dynamics can only be observed in the case of very low FP values, and even there such dynamics do not always turn up. This shows first that FP is probably still too rough to allow more detailed prognoses; and second, it shows that a single parameter can (possibly) never be sufficient for the generation of complex dynamics, which already goes for deterministic systems (see below, section 3.1.2.). However, at this point we can already formulate a very general conclusion regarding the correlation between deterministic rules and additional stochastics:

---

[41] Unfortunately, the mapping of a rule system into the according frequency matrix is only injective and not bijective, because different rule systems can be represented by the same matrix.

If the deterministic rule system generates only simple dynamics (Wolfram classes I and II), then the introduction of probability parameters only rarely leads to more complex deterministic dynamics. On the other hand, complex deterministic dynamics can be simplified quite easily by probability parameters. *It seems to be much easier to increase than to decrease order.* This is already a non-trivial result *per se*.

As I already mentioned, social systems practically always have to be characterized by some sort of duality of deterministics and stochastics. Therefore, these initial results, however preliminary they may be, already provide some not insignificant insights: even the continual deviation from social rules does not lead into chaos (in the mathematical sense), instead, it simplifies dynamics that are *per se* complex. There seems to be a lot of research left to be done about this issue.

Finally, a special advantage of system analysis realized by CA and related modelings stems from the following fact: those modelings can be much more directly related to concrete experience than would be possible using traditional mathematical techniques. The point of this claim can be explained by the fox and goose system:

In the tradition of mathematical ecology, modelings of predator-prey systems usually have the form of the well-known Lotka-Volterra equations, where the covariation of the predator and prey populations are expressed by two differential equations. Yet populations are abstract concepts. What can really be experienced by observing actual foxes and geese is that foxes eat geese, geese eat plants, and both species have a certain number of descendants per year; moreover, both species have a certain "natural" life span. In contrast to the Lotka-Volterra equations, a modeling of this predator-prey system with a CA directly begins with these experiences and transforms them into the model. As a further contrast to the equations above, it is taken into account here that the prey has to be within geometric reach for the predator; the mere fact that there is prey at all is therefore insufficient. This claim also corresponds to concrete observations of real predator-prey systems. Of course, such concrete experiences must have been systematized in the sense that the parameters of the interactions must have been empirically determined. However, it is not necessary to take the system as a whole as the level of abstraction, as in the case of the equations of Lotka and Volterra; instead, one can go from the level of empirical observations directly to the level of the formal model. Gerhard and Schuster (loc. cit.) have described this advantage of the CA compared with traditional mathematical techniques very vividly with regard to chemical systems (see also Toffoli 1997); this advantage is even more important for the modeling of social systems,

since there traditional mathematical modeling techniques can be used only in very few cases, and even these have to be reduced to a great extent. I have repeatedly pointed toward this fact already.[42] With CAs, however, it is possible to begin directly with the observation of social interactions and construct a formal system model. With this procedure the interpretative social research as a part of a systematic microsociology can gain a new status too.

The so-called "Boolean nets" (BNs) (or "genetic nets") which were used in particular by Kauffman (see above, section 2.7.) form a very similar class of "pure systems" that can be assigned to the paradigm of self-organization. Since these offer very instructive parallels and extensions to the results that we just described for the CA, we will give a brief analysis of this system type here too.

3.1.2. Boolean Networks (BNs): Making (Propositional) Logic Dynamic

The basic idea of the BN was already developed in 1943 by Mc Culloch and Pitts (loc. cit.); as "logical switchnets" they are known to be the logical basis of computer hardware, and their principle has also entered portions of the logic of artificial neural nets. As was outlined in section 2.7, BNs are as simple as CAs. In fact, they follow very related principles:

We take as a starting point the logical bonds of the propositional calculus that combine artificial units; similar to the cells of a CA and the artificial neurons, these logical bonds can take different states - BNs in the strict sense can take two. The classical logical bonds of the propositional calculus are the logical relations "AND" (conjunction), "OR" (disjunction), "IF THEN" (implication), "IF AND ONLY IF" (equivalence); etc. In the simplest case, these logical bonds combine two propositions; we can then speak of logical bonds or operators with two variables. The negation "NOT" is an exception: it transforms one proposition into another one (namely the negated original proposition) and is therefore a logical bond with only one variable.

Since the propositional calculus is formal, logical bonds are formally defined, namely by so-called truth tables or truth matrices. For example, a disjunction is only false if both propositions are false, else it is true; in contrast, a conjunction is only true if both propositions are true, and is false in all other cases. The implication a $\rightarrow$ b is only false if a is true and b is false; else it is true. Finally, the equivalence a $\leftrightarrow$ b is true if and only

[42] Gerhard/Schuster also show that CAs are often superior to the established techniques of modeling with differential equations as regards their "computational facility": they make it possible to compute system dynamics in "realistic" computer time where even very powerful computers would need too long to compute the respective differential equations.

if a and b have the same truth value, i.e., if they are either both true or both false; in both other cases the equivalence is false.[43] The other logical bonds are treated accordingly, and it should be clear already that there are exactly $2^4 = 16$ logical bonds, or Boolean functions (BFs), when there are two variables. In general, there are exactly 2^{2^n} BFs with n variables, which can be proved just as in the case of the number of rules of CAs.

Thus, a BF with n variables is defined by the fact that one specific truth value is assigned to each of n combinations of truth values. If we abstract from the fact that the elements of these systems are propositions and the BFs are the colloquially known logical bonds of propositional calculus, we obtain the possibility of constructing artificial systems whose elements can be in one of two states - 0 or 1, off or on, dead or alive, false or true - and these states result from the respective BF that combines an element a with n other elements. If $n=2$ and the respective BF is a disjunction, then the element a is in the state 0 if and only if both elements b and c that are combined with it - "switched together" - are in the state 0. If one or both elements are in the state 1, then a is in the state 1. This is exactly the logical principle on which the "logical lattices" of computer hardware work. (Conventionally, the states 0 and 1 correspond to the propositional terms "false" and "true," respectively.)

As outlined in section 2.7, Kauffman (1992, 1993, and 1995) transferred this constructional principle to the development of artificial nets; I have already mentioned the biological motivations for that. These BNs can be considered as formal dynamic systems in exactly the same way as CAs, and can be analyzed accordingly. The rules of interaction between the elements are of course the BFs that are introduced for a certain network, and the state of a net is the totality of states that its elements have attained at a particular time. Again, the dynamics of the system results from the fact that from a specific initial state onward, the single elements reciprocally influence each other according to the BFs; if we describe the totality of BFs that are determined for a network as the system function or transition function f, then the system dynamics exhibits itself as the iteration of the system function (practically always in a non-linear way). Likewise, we can classify these dynamics according to the Wolfram classes and ask for the control parameters that generate specific classes of complexity.

---

[43] This procedure of defining propositional logical bonds by truth matrices is also the reason for the fact that the propositional calculus - in contrast to the predicate calculus - is a decidable theory. Any formula of the propositional calculus is made up of variables for the propositions and different logical bonds, which makes it possible to check its truth immediately by use of the respective matrix.

Indeed, it became clear rather quickly that the BNs' dynamics can really be divided into our four classes, and that very specific parameters are responsible for that. The first parameter, discovered by Kauffman, is the number of elements that have an effect on one element; in terms of the propositional calculus, this is the number of variables of which a logical bond consists - the generally accepted expression for that is K, according to Kauffman. In the above example, it is obviously $K=2$. One can roughly summarize that the K parameter means that if $K = 1$ there is a network behavior of Wolfram class I; for $K = 2$ - as it was already mentioned in 2.7. - a dynamics of class II arises; and in the case of relatively extensive networks with a number of elements of $N > 10^4$, we also have class IV.[44] This class also arises if $K=3$. If $K > 3$ the dynamics becomes chaotic, that is, the system turns into class III.[45]

A parameter that is independent of K was found by Weissbuch and Derrida (Gelfand/Walker 1984). It is the "inner homogeneity" P of a BN (Kauffman 1993), where P is simply defined as the deviation from a homogeneous equal distribution of 1 and 0 in the results of a BF. Logical implication, for example, has a P value of 0.75, because its value table is 1,1,1,0. P obviously varies between 0.5 and 1. If P is to be defined not just for one single BF but for a whole network, the P value of the network is of course the mean value of the network functions' P values.

For P - understood as the P value of networks - it is valid that $P = 1$ yields the most ordered state, namely a dynamics of Wolfram class I. If P is gradually decreased, classes II and IV arise; if P approaches 0.5, the network turns into a chaotic dynamics, that is, it becomes a class III net. This is not surprising if we consider its obvious similarity to Langton's λ parameter: in both cases the system dynamics depends on how homogeneous the possible states can be reached by the rules. If this is possible in a homogeneous way, the system is largely chaotic; it will take very many system states in succession and it will then no longer be possible to predict its behavior. If the system's transition rules are strongly restricted and particular states are significantly preferred, then P and λ work as control parameters, that is, they force the system into relatively simple trajectories with simple attractors.[46] There seem to be "universal" order patterns at work at this point.

Kauffman (1993, 203) has developed another parameter which he described with the concept of a "canalizing Boolean function". These

---

[44] Since K and N are established expressions, one often finds the term NK system in literature.

[45] As will be shown below, K is actually not an independent control parameter, because its values depend on the values of other parameters.

[46] Both Kauffman (1993, 219) and Langton (1992, 87) have noticed the similarity of P and λ.

functions are BFs for which it is sufficient to know the value of *one* variable in order to know the result of the function (in terms of the propositional calculus: to know the proposition's truth value). For example, the disjunction a v b is canalizing because the truth of the whole proposition follows from the truth of a; accordingly, the same may be said for the truth of b. Thus, in case of a network with canalizing functions, we only need to know one unit to know the effect of the function.[47] XOR (that is, the excluding OR) is, for example, not a canalizing function.

As for the canalizing functions' "ordering" effect, we can say that the more of them there are within a network, the more ordered are its dynamics. Kauffman (1993, pp. 203) presents some statistical results whose details can be neglected here; the relation V of canalizing functions to N (the number of units) seems to indicate a critical limit. If V = 1 or higher, the network dynamics belongs to class I, II, and sometimes IV respectively.[48]

Jörn Schmidt and I have examined another control parameter for BNs for the case K = 2. Taking up a suggestion by Kauffman, we were interested in the "topology" of Boolean nets, that is, the special geometrical forms of connecting the units. In contrast to K, which defines how many units *have an effect* on another unit, we analyzed what consequences the different values have on *the number of units that one particular unit has an effect on.* We found that the respective values also form a control parameter. Analogous to the P parameter, we took the single units' mean value difference as the degree for this parameter v.[49] If v is very small (that is, if all units have an effect on about the same number of other units), then the system develops a dynamics with a chaotic tendency and very long periods; the greater v is (i.e. the greater the possibilities of units' effects differing), the simpler are the trajectories - these systems belong to Wolfram classes I and II.

Sociologically speaking, v can be interpreted as a "degree of democracy": if democracy is not defined merely with regard to the formal equality of the members of a society but also from the equality of their practical possibilities of acting effectively, then democratic societies can easily be

---

[47] Of course, this does not apply to every value of a canalizing function's units: Knowing that a has the value 0 in a disjunction does not say anything about the function's effect - if b = 0, the effect is 0; if b = 1, the effect is 1. If there is at least one variable that has a value to which the above condition applies, the function is canalizing.

[48] Kauffman (1993, 204) points to the fact that practically all "regulatory genes" (see above, 2.7) have their effect because of canalizing functions; this could explain the high degree of order inside a genome.

[49] The exact technical definition results from a specific kind of matrix representation of the BN, and can be neglected here.

modeled as BNs with a very small v; the reverse goes for the opposite case, that is, for non-democratic societies. This understanding of democracy, which goes beyond formal equality, has always been used as a critique of merely formal democratic societies, and has a venerable tradition; Anatole France has expressed this position most concisely in a famous aphorism: the law, in its majestic equality, forbids the rich, as well as the poor, to sleep under bridges.

How are these parameters connected? Up to now this question has not been systematically investigated; Kauffman often treats the parameters as independent, and he laconically remarks in this context: "Undoubtedly, further control parameters will be discovered" (1993, 219). However, an initial connection can be shown between the size of K, that is, the number of variables of the Boolean functions, and the number of canalizing functions. On the one hand, there are a high number of canalizing functions for K = 2: only two of the 16 functions are not canalizing (namely the XOR function and equivalence). Thus, the proportion of canalizing functions is 0.875 for K=2. For the 64000 functions of K=4 the number of canalizing functions is only 0.05, and this number decreases with an increase of K. Hence, there is a very small probability that an arbitrarily formed network with K > 2 will consist mostly of canalizing functions; at the same time, this makes the high degree of order for K = 2 plausible (Kauffman 1993). K is just a derivative parameter, as can be easily proved by experiment. It is quite simple to generate complex dynamics for K = 2: one only has to reduce the number of canalizing functions.

Furthermore, there is also a connection with the P parameter, because canalizing functions obviously promote an unequal distribution of 1 and 0. For example, if K = 2, two of the 14 canalizing functions have the P value 1, eight have the P value 0.75, and only four have the P value 0.5; the two non-canalizing functions have P = 0.5. For K > 2, the number of functions with high P values rapidly drops. But these are only statistical values; therefore we can say that the single parameters can be varied independently and that - aside from K - they are also logically independent. Since all known parameters - including v - allow exceptions, one single parameter cannot be a sufficient condition of the prognosis of special system dynamics. Jörn Schmidt and I have investigated combinations of different parameters and, for example, for the combination (P,v,CF) we found regions where these combinations were sufficient - although this is not even true for all regions (CF is the proportion of canalizing functions to the total number of functions). In other regions, this combination is just necessary (whereas the single parameters for themselves are not). To be a bit more precise: apparently, the different parameters are sufficient in the case of simple dynamics (class I and II) but not necessary. For complex dynamics a single parameter is necessary but not sufficient;

only particular combinations of several parameters are sufficient in most cases.

I have already noted the close relationship between parameters λ and P. This relationship results mainly from the fact that, in general, BNs and CAs are quite similar formal systems; Wuensche (1994) even describes CAs as special cases of the general BNs. This becomes evident by admitting more than two state values for the units, which implies departing from the logical model of propositional calculus. In this case, the units that directly affect other units through a BF can indeed be conceived as a "pseudo neighborhood" of the dependent unit. But in contrast to typical CAs, the rules for these pseudo neighborhoods can locally vary to a high degree; CAs usually have the same rules for different particular combinations of cell states. Removing this limitation for CAs as well (a move for which there is a lot to be said from a sociological point of view) would lead in practice to a logical equivalence between these two formal models of self-organization; which of the two techniques to use would therefore depend solely on the specific research interest or simulation practice in question. Accordingly, the rule systems of CAs could also be characterized with CF and v.

With the concept of canalizing functions we can also understand the connection between the parameters Z and λ within CAs. We know that Z measures the probability of cell states being clearly determined in the previous states of the automaton. As a consequence, if a function f is canalizing, Z is necessarily relatively small; and if there is a large number of canalizing functions, the system has a high degree of order. It is easy to see that in the case of low Z values the system function as a whole has to contain a lot of canalizing functions - transition rules for the cells - and vice versa. In accordance with the canalizing functions' relevance for the values of the P parameter (with which we have already dealt) there is a corresponding relevance for λ, which results in a kind of "natural" correlation between the parameters Z and λ, as it were. In particular, it turns out that λ not only measures a system's "openness" towards changes (of the cell states), but - because of its connection with Z - also says something about the system's "temporality".

Yet recent investigations about the topology of BNs by Jörn Schmidt and me (Klüver and Schmidt forthcoming) have shown even more and deeper connections between the different control parameters. I already mentioned the "topological" parameter v which demonstrates clearly that the topology of BNs (in CAs the topology is usually only homogeneous) has an important impact on their dynamics. The results concerning v can be generalized. Another topological parameter is, for example, the average distance d between two elements of a BN: if the mean distance

between two elements is nearly the same in a BN, then its dynamics is rather complex and vice versa - one immediately sees the parallel to v.[50] Therefore the following conjectures from our results and the results mentioned above are rather obvious:

(a) In formal and social systems one has to distinguish between two kinds of rules. The first kind may be called "general rules" because they define the interactions between the elements of the system in general, regardless whether interactions between particular elements are possible at all. The general rules can be measured and classified by control parameters like P or CF; we saw that it is possible to a certain degree to understand - and explain - the type of dynamics the rules of one system may generate by knowing the values of these parameters.

The second kind of rules determines which interactions are possible at all, logically in the same way the geometry of the physical space allows or forbids particular local interactions of elements of physical systems. Therefore, and because these rules can be analyzed in a geometrical way, e.g. by using graph theoretical methods (see below, section 4.4. and Klüver/Schmidt loc. cit.), we may call these rules "geometrical" or "topological rules". The geometrical rules represent those characteristics of a system which are usually called the "structure" of a system (e.g. Freeman 1989) and which have been analyzed rather thoroughly in social network analysis. Yet this structure has an important impact on the dynamics of the system too, which has been overlooked by mere structurally oriented researchers. Our researches have shown that the geometrical rules of a system, that is its structure, are probably more important for the dynamics of a system than the general rules.

(b) By looking a bit closer at the different control parameters of both general and geometrical rules one common characteristic seems rather obvious: all control parameters - with the exception of K - measure different degrees of equality or inequality respectively. P, for example, measures the equality or inequality of reaching particular states by general rules; canalizing functions can be understood as decision rules where just one member decides on the outcome of the decision process; therefore the number of canalizing functions in a system is a measure about the equality of influence in decision processes.[51] The stochastic parameter FP (see above, section 3.1.1.) is obviously a measure similar to

---

[50] To be a bit more precise: this parameter is measured as the variance var(1/d) of the inverse distance d between two elements.

[51] A well known example for a decision process which can be characterized by a canalizing function is the voting procedure in the UN security council: as each of the five permanent members has a veto right, they can be represented by a logical conjunction with K = 5 which is a canalizing function. It is obvious that this procedure divides the members of the security council in two classes of inequality.

P as it determines the probability of reaching certain states or not. Remember that low values of FP, that is equal distribution of the frequency and probability values, tend to yield complex dynamics and vice versa.[52]

Even more plain is the measure of inequality in the case of the geometrical parameters, that is the control parameters of the geometrical rules. v has been examined in that respect before; the distance parameter d measures whether a system is connected metrically rather homogeneously or whether there are for example subsets of elements only connected with the rest of the elements by chains of other elements. The latter case represents for example social systems which contain strict hierarchies, that is, some elements can interact immediately but with others only through a lot of intermediaries. Obviously, d is a measure of inequality; if we model social systems with BNs with high or low degrees of d respectively we get measures of the hierarchical structural properties of the system.

This being the case it seems safe to postulate a "conjecture of inequality":
The different control parameters all measure different degrees of equality or inequality respectively of a system. Since there are always different dimensions of equality or inequality in each complex system, there are necessarily different control parameters also. In social systems for example inequality is represented by differences in status, income, sex, education and so forth. Now, if a system has rather high degrees of inequality in more than one dimension, its dynamics is always simple; the more aspects of equality a system exhibits, that is, the more the rules of the system can be characterized by values of the control parameters characteristic for formal equality, the more complex the dynamics of the system will be.

This conjecture is possibly able to explain some of the differences between modern democratic societies and traditional ones: in the first case we have systems which contain rather high degrees of equality; accordingly they can be modeled by formal systems whose rules can be characterized by parameter values characteristic for complex dynamics; complex dynamics are obviously a feature of democratic societies. On the other hand, traditional societies with high degrees of inequality

---

[52] The only exception is K which is defined as an absolute value and not as a proportional value like all other control parameters. Therefore and because we (Jörn Schmidt and I) could not confirm some of Kauffman's results in respect to K, we have some doubts whether K is an important parameter at all. But these are problems for specialists.

exhibit only simple dynamics - their parameter values are characteristic for the latter.

Yet, as we shall see more clearly in chapter 4, the complex dynamics of democratic societies, due to their high degrees of equality, produce another consequence which may be called the immanent coercion to reform. Since the dynamics of democratic societies are rather complex, they inevitably change their states rather often and therefore deviate from favorable states they achieved in the immediate past. In order to correct these unfavorable developments the democratic systems have to change their rules to reach favorable states again. But since the new rules must preserve the former degrees of equality (else the systems would not be democratic anymore) the dynamics generate necessarily unfavorable states again. So the rules must be changed again and again - equality obviously has its price. Who rides on a tiger cannot get off.

CAs and BNs are *the* formal models of systemic self-organization, and we have seen that all system dynamics of interest can be generated with them. In particular, it seems possible to make sure of fundamental connections between the different classes of system dynamics and the actual rule systems. Naturally, despite the partially amazing parallels to biological systems (which have been particularly emphasized by Kauffman), we can not derive the behavior of real systems *directly* from the above results, because it has first to be shown that it is possible (and feasible) to map real systems into those formal systems in a model. Nevertheless, with all due caution, the considerations just made – especially with regard to social systems – seem rather safe.

Obviously, we do not get anything for free here. Systems that are completely ordered (e.g. class I systems) are also the least efficient - for example, when it comes to information processing.[53] The previously mentioned metaphor of the "edge of chaos" shows that systems that are "just within" the bounds of order but already exhibit quite complex dynamics are the most efficient ones, which makes it most difficult to understand them. The "*Unübersichtlichkeit*" of modern societies (difficulty in understanding; Habermas), which is often stressed in present social theory, may also be due to the following: the efficiency of modern societies is closely connected to this fact. Premodern societies, which belong to class I and II, are easier to comprehend because their respective parameters have different values; accordingly, they are less efficient.

In chapter 2., we repeatedly emphasized that social systems are adaptive; this means in particular that they are able to change their rules, and thereby their control parameters (see below, section 3.3.). Thus, the

[53] Sometimes, the systems' ability of processing information is plainly seen as the degree of their evolutionary level (Bedau/Packard 1992).

control parameters do not function as constants there; as in the case of certain linguistic universals (Pinker 1994), they rather have the character of "implicative" universals: a social system can obtain a rule system with particular parameter values by adapting to its environment. Thereafter, however, it is left to these parameter values *without further rule changing*; the dynamics that results from the rules entirely depends on the actual parameter values. Therefore, it is certainly no exaggeration if I speak of "universals of general rule systems" in the context of the control parameters.

Finally, with these pure systems we can understand the fruitfulness of the concept of attractors in a more concrete way. We have characterized the different system dynamics as outcomes of the trajectories' different attractors. Hence, if we know a system's attractors, we know its general system dynamics. In particular, the fact that a system reaches a specific attractor can be interpreted as the output of a specific system performance (which reaching this state) and the subsequent stabilization of the system at or near this state of performance. Allow me to quote a longer passage from Kauffman (1993, 191) that clearly shows how general this thought has been developed in the theories of different dynamic systems:

> "Dynamical systems ranging from genomic systems to immune systems, neural networks, organ systems, communities and ecosystems all exhibit attractors. The alternative attractors in a genomic regulatory network ... can be interpreted as the alternative cell types in the organism ...; the alternative attractors in immune networks correspond to different immune states ...; the alternative attractors in neural networks have been interpreted as alternative memories or categories by which the network "knows" its world ...; the alternative attractors in cardiac systems correspond to normal and abnormal rhythms ...; the alternative attractors in ecosystems correspond to alternative stable patterns of species abundance."

This list could easily be extended, especially by sociological applications: different attractors within a social system that are realized simultaneously correspond to different social subsystems, and different attractors within socio-cultural evolution correspond to different types of societies (such as tribal societies, state-organized societies, etc.) (cf. for example, Yoffee 1994).

The fact that different trajectories can lead to the same attractor led us to an extension of the attractor concept (already mentioned above), namely an extension towards the attractor's "basin of attraction" (Wuensche 1994), that is, the number of all trajectories that can lead to that attractor. Kauffman (1995) clarifies this concept with the beautiful picture of different rivers flowing into the same sea.

If we add yet a further extension, namely the concept of the basin of attraction field (Wuensche/Lesser loc. cit.) - meaning the number of a system's basins of attraction and their topological interrelations -, then we can describe a dynamic system mainly by this "geometry" of its state space.

In the 2nd chapter I repeatedly pointed to the fact that though the concept of self-organization is necessary for the analysis of systems of the cultural and cognitive sciences, it is not sufficient in itself, because it does not take into account the fact that the systems that are of special interest here are capable of systematic rule changing. Both CAs and BNs change their rule systems only through external operations. But if we want to grasp adaptation within pure systems as a *characteristic* of systems, we will have to go beyond the systems we have examined so far and construct rules as well as meta-rules. This can be achieved in different ways; for various reasons, we at Essen experiment primarily with combinations of single rule systems with GAs. That is why we are going to take a closer look at these in the following.

3.2. GENETIC ALGORITHMS (GAS): SELF-ORGANIZATION THROUGH ADAPTATION

> " ... it is easy to fall into the trap of perceiving GAs *themselves* as optimization algorithms and then being surprised and/or disappointed when they fail to find an "obvious" optimum in a particular search space. My suggestion for avoiding this perceptual trap is to think of GAs as (highly idealized) simulation of a natural process and as such they embody the goals and purposes (if any) of that natural process. I am not sure if anyone is up to the task of defining the goals and purpose of evolutionary systems ..." (De Jong 1985; quoted after Michalewicz 1994, 15)

The subject of "evolution" and, by the same token, of "adaptation," makes many people besides sociologists (and biologists) think of received evolutionary biological theories. Therefore, it is almost natural that sociologists as well as mathematicians and computer scientists were influenced by biological models when they - each in their own specific way - tried to grasp evolutionary and adaptive processes in their own sciences.

We are going to take up the problem of sociological evolution theories again in the course of this book (cf. e.g. Giesen 1980). If only because of its semantic bearings, the frequent use of biological concepts in the computer sciences is conspicuous; in particular, this goes for programs that are supposed to perform optimizations as adaptations of artificial

systems towards given goal values (in one way or another). I have already outlined an orientation to neurophysiology and learning theories based thereupon in my remarks on artificial neural nets; an evolutionary biological terminology is linked to concepts such as "evolutionary strategies" (Rechenberg 1973), "genetic programming" (Koza 1992), "evolutionary programs" (Michalewicz loc. cit.), and the previously mentioned "genetic algorithms" (GAs).

Of course, there had been a tradition of mathematical optimization algorithms long before the corresponding biological orientation; special optimization techniques have been developed recently (e.g. in AI research) like "hill climbing" and the physically oriented "simulated annealing" (Davis 1987). However, the explicit alignment with biological mechanisms is relatively new; therefore, it is not surprising that these "biologistic" procedures show some specific strengths and weaknesses, which they partly share - see De Jong above - with their models. If only because of that, it seems reasonable to conceive of GAs (with which we will deal now), as formal systems *sui generis* just as we did with CAs and BNs in section 3.1. And yet, De Jong's note also shows that this point of view is by no means self-evident, and that the main part of the "GA-community" (Michalewicz) without hesitation locates GAs in the tradition of optimization algorithms.[54]

In the following, I will always use the term GA with its classical meaning, as developed by Holland (1975). Variants, like e.g. the elitistic GA, will be explicitly named as such. The principles and possible sociological uses of GAs have already been described in Klüver 1995; I would like to outline the main points once again:

GAs can compare strings of symbols according to specific criteria, such as: pick out the "best" of each; subsequently recombine these best strings ("crossover"); vary some of the strings' symbols at random ("mutation"); and compare the newly made strings again with regard to their "fitness". Thereafter, the best are again selected and subjected to the procedures of crossover and mutation. There is then a new selection of the best according to criteria of fitness, and so on, until a desired goal - a global or local optimum - is reached. It is obvious that this is a simulation of the biological evolutionary mechanisms of heterosexual reproduction and mutation incorporating genome variation, hence the name "genetic".

---

[54] The literature on GAs, which is hardly manageable in the meantime - there are more than 300 titles in Michalewicz alone -, almost exclusively deals with technical problems, e.g., how special problem classes can be handled in a better way through modifications of the "standard GA". I will not go into that here; instead, I refer to Michalewicz' instructive book for a general survey.

There are various possible applications of this special programming technique, which can be seen from numerous reports (e.g. Holland 1975; Goldberg 1989; Schöneburg et al. 1994; Michalewicz 1994); GAs seem to be particularly suitable for the optimization of other programs, with which they are combined to form "hybrid" whole programs (Klüver 1996; Hillis 1992; Koza 1992; Belew et al. 1992). I will go into that in more detail in section 3.3., so this note will suffice here.

Thus, if we conceive GAs as systems *sui generis*, we have to look for the "natural" analogy in evolutionary biological "systems" rather than in the field of physical-chemical dynamic systems that comprise the fundamental point of reference for CAs and BNs. However, GAs are not necessarily meant to copy natural gene variations and selections in all their subtleties; the natural processes serve instead as heuristic guides toward which the various patterns of GAs may orient themselves. So, again we are not dealing primarily here with simulations of particular "real" systems (which are of course also a subject of research), but instead mainly with "pure" systems, which serve particularly well for the analysis of specific forms of system behavior.[55]

We already mentioned that the "elements" of a GA consist of strings of symbols that are described as vectors. Let there be n strings consisting of k symbols. In hybrid programs the symbols are nothing but operational instructions on whether (and if so, how) specific rules of the related subprogram are to be run. Furthermore, let there be a given "evaluation function" which assigns specific values - in the form of a scalar - to each string of symbols. Thereby the evaluation function defines an ordering relation for the elements; it may of course occur that two elements get the same value.

A transition function for our system is now defined by the following procedure:

1. Order the elements according to the evaluation function by decreasing values.
2. Select the m best of n elements, where m = n/2 approximately (this is an empirical value).
3. Remove r arbitrarily selected components from the best element, where $r < k/2$.
4. Remove r components from the second best element following the same procedure; exchange these components from the best and the second best element (crossover).
5. Apply the crossover principle with the best element to all remaining elements, and subsequently proceed with the other elements in turn until n new elements have developed.

---

[55] So the above quotation from De Jong must not be misinterpreted in a naturalistically reduced way, and to my knowledge that is what De Jong intended too. Not for nothing, Holland distinguishes between natural and artificial systems already in the title of his book (1975).

6. Change individual arbitrary components of the new elements with a specific rate of probability (mutation).
7. Order the n new elements according to the evaluation function.

This obviously defines specific rules of our elements' interactions with each other: the elements interact with each other on the analogy of heterosexual reproduction (including variations of the newly developing "genome" caused by mutation), and thereby they yield new elements. In order to rid oneself of this biological terminology, one might also understand these rules on analogy with the way that a specific string of symbols describes the state of an element (similar to the CAs): as a consequence of an interaction with another element, i.e., an element in a different or the same state, the first element turns into a new state or it remains in the same state (which cannot be ruled out in the case of our stochastic systems). (This especially occurs when our system has become "evolutionarily stuck"; in this case, the vectors - the elements - do not change, at least for the short term.) This clearly determines, first of all, how new elements arise from previous ones or how their state values change. This results from the fact that each newly generated element is simply assigned to the corresponding "old" element and takes the same place in the new order that the old element had in the previous one. Second, this determines a transition from one system state to the next, which simply consists of the new elements, their order, and, of course, the values of the elements. It has to be mentioned, though, that the sequence of system states of GAs is not strictly recursive: this follows from the fact that both crossover and mutation are subject to random processes. In fact, they are subject to such processes with every single recombination of two cells; in this respect, we have, in principle, a stochastic system.

If we continually repeat the above procedure we obviously obtain an iteration of the transition function again, as well as non-linear dynamics. In fact, the evaluation function does not necessarily produce better values for the new elements, in which case it would be possible to describe its growth graphically as the permanent increase of a straight line. The values rather decrease almost always for some time, or at best they remain constant. Nevertheless, for the long term one can be sure that the values will tend towards a certain optimum, regardless of short-term drops. How can that be explained?

At first glance, it may seem intuitively plausible that a crossing of the best elements yields better and better results. When one considers this matter again, however, it is not plausible at all for the following reasons:

System development usually begins with elements that have no particular quality regarding their evaluation function. The initial elements

are often even constructed at random; in the case of large strings of symbols, there is not much chance of having highly evaluated elements at all - and after all, what use would a GA have if one already knew from the outset what excellent elements would have to look like? A crossing of elements with medium (or lesser) status does not necessarily produce outstanding results. Neither is it the case that mutations are a guarantee for success, since we know from genetics that the outcome of mutations of individual genes is for the most part negative. This is equally true for the GA's strings of symbols - as far as mutation is concerned, success is always the improbable case. To make things even worse, GAs are "blind" like the evolutionary clock-maker in Dawkins (1987): With each new "generation" of elements all the elements of the predecessor generation vanish, including the best such elements - even if the best members of the predecessor generation might actually be superior to their descendants. Thus, the GAs "forget" their positive experiences (the negative ones are forgotten anyway, since they are already sorted out in the 2nd step of the transition function).[56]

In view of these problems it is not surprising that the GA was at first received with irritation and skepticism (Levy 1993, 153). The obvious successes that could be achieved with it indicated something that actually seemed impossible, namely "the essence of making something out of nothing" (Holland, quoted after Levy loc. cit., 159). In the meantime some mathematical proofs (and innumerable experimental results) have been provided to support the fact that the GA almost always will improve its elements, unless there happen to be optimal elements present at the beginning. A proof of this that is particularly elegant in my opinion can be found in Michalewicz (1994, 65), employing a famous theorem from metrical topology. I will briefly outline the gist of that proof, since it is an excellent example of how pure mathematics, computer theory, general systems theory, and, hopefully, also sociological systems theory (can) at certain points enter into an intellectual symbiosis:

Metrical topology is concerned with the properties of so-called metrical spaces, that is, sets for which it is possible to define distance relations between the elements. Ordinary length measuring is a special case of those distance relations or metrical relations. More than seventy years ago, the Polish mathematician Banach put forward and proved the so-called Banach fixed point theorem for particular metrical spaces. It basically says that in specific metrical spaces - so-called complete spaces

[56] There are exceptions, namely the so-called elitistic GAs, to which I will come back later; there, the best element of one generation is usually kept and passed on to the next generation without being changed. This is different, by the way, in the so-called evolutionary strategies of Rechenberg (1973), where the old elements vanish only if they are all worse than the selectable new ones.

- every "contractive" function has a fixed point towards which the function is, so to speak, bound to move. A contractive function is defined by the fact that the distances between two function values diminish the more often the function is applied to these values - the function contracts the elements of the metrical space. In other words: if x and y are two elements of the metrical space, f is a contractive function, and d(x,y) is the distance between the elements, then

$$d(f(x),f(y)) <(=) d(x,y) \text{ and } d(f^n(x),f^n(y)) < (=) d(f^{n-1}(x),f^{n-1}(y)).$$

Again, f^n is the n times iteration of f, that is, the n times application of the function to itself. If this iteration is carried out a sufficient number of times, then, according to Banach's theorem, the function f necessarily converges to the fixed point x_0 of the function f, regardless of the initial value.

As observant readers, you can probably guess what will happen next. We take all the strings of symbols on which we want the GA to work (i.e. the elements of our formal system) as a metrical space. We can obtain a distance definition from the evaluation of the elements and show that this is a complete space (I will leave out the technical details here). We then define a contractive function, which is of course just the slightly modified transition function of steps 1 to 7.

We can easily prove that this function is contractive (the decisive step of the proof, after all), and it immediately follows from the Banach fixed point theorem that there is a fixed point for the GA - a specific optimal value of the evaluation function - toward which our system converges. So the convergence of the GA is simply a consequence of the "contractive property" of its transition function.

It really does not matter with which strings of symbols we begin: the above mentioned blindness of the GA and the usually quite faulty initial set of elements may be able to prolong the process, but they cannot decisively impair it. It is, however, required that the transition function is iterated a sufficient number of times, which - in the case of very large strings of symbols - might lead in practice to a considerable amount of computer time. Furthermore, all this does not mean that the fixed point is exactly the value that was desired when the program was constructed. The *attainable optimum - the fixed point - might still be suboptimal for*

the user, and the program may therefore require further measures (see, e.g., Hillis 1992).[57]

The previously described case concerns convergent system behavior, that is, the development of the system states towards a specific state. So, in a way, we are dealing with a system approaching states that will make it a system of Wolfram class I or, more likely, class II, because the convergence point is never precisely reached after a finite number of steps, and the system "oscillates" somewhere near that point. Thus, the convergence point or fixed point determines a basin of attraction for our system, which it may, however, leave again in the long-term. This mainly results from mutations, which can change the elements in such a way that the system is virtually forced to direct its trajectory to new basins of the state space.

Strictly speaking, a GA is not *per se* an adaptive system as defined in section 2.2., since the rules of interaction - genetic operators and selection - do not change. "Only" the values of the single elements change. Of course, we can vary these rules themselves; Michalewicz (loc. cit.) lists different variants of genetic operators, and one can construct GAs that change their own rules dynamically. In the classical sense, however, we can speak of an adaptive system with a GA only if the GA operates upon another system and regularly varies the rules of the latter through its transition function. I already mentioned that those programs are called "hybrid" programs, and such programs will be the subject of our next subchapter. However, in order to clarify and to explain the relevance of these programs for systems theoretical analysis, I will now illustrate them with the example of biological evolution.

Descriptions of elements as "chromosomes" and elements' components as "genes", which can be found in the literature on GAs, often fail to mention that there is an important difference between GAs and biological organisms, namely with regard to the difference between genotype and phenotype. In a pure GA, genotype and phenotype are

[57] This seems to bear a slight contradiction: a contractive function is "monotonous", that is, it only develops in one direction - shrinking of the distances. On the other hand, a "normal" GA can also "go back", that is, its values can decrease and the distance to the optimum can grow. In fact, a standardized GA, as described above, is not contractive in a strictly mathematical sense; therefore, Michalewicz modifies the above transition function so that it becomes contractive. The normal GA is, so to speak, only globally contractive, that is, in the case of sufficiently many iterations of the transition function there always occurs a convergent behavior. Michalewicz himself points to the fact (loc. cit. 69) that the proof strictly applies only to infinite populations; in practice it can always happen that the program gets stuck in suboptimal populations - elements - and cannot get out of the suboptimum. This frequently occurred in our experiments. Mathematically speaking, this means that Michalewicz' proof is not "constructive": there is always a fixed point, but we do not know how the GA gets there and if that can happen in finite time.

identical, as it were, which means that the selection directly effects the genotype. In biological systems things are different of course. There, selection only effects the phenotype, while genetic operators work on the genome. Therefore, the evolution of organisms and species can be described in systems theory terms as the dynamics of hybrid systems; this becomes clear in the frequently discussed model of Kauffman, which has some important empirical facts supporting it:

Let us think of a genome as a genetic BN whose dynamics can be discerned from its single BFs. This network can have one or more attractors; in the latter case, the single attractors correspond to the different types of cells. Now, if several of these - similar - genomes are exposed to the genetic operators of mutation and crossover, the BFs are thereby varied and the new genome realizes different attractor states. Consequently, this leads to new cell types, which either do or do not prove worthwhile in the environment – according to the evaluation function. If the genome is furthermore capable of autocatalysis (that is, if it can produce new genes), then new genes will be integrated into the net and coupled to the old genes by BFs, by which process new and more numerous cell types are produced. *Whether* this happens or not depends of course on the selection, that is, on the value of the newly developed phenotype. Thus, the "natural" GA operates on the rules of the genetic system BN so that the system and GA together can be regarded as a hybrid system.[58]

If the genome in question is in an attractor state and is "disturbed" by the GA, then its further development depends on both the degree of the disturbance and the complexity class of the genetic network. If the net, e.g., is in class II, then the system will probably return to its previous attractor if the disturbance is not too strong - in other words, if the new trajectory of the genetic net belongs to the attractor's basin of attraction. That is why species that are largely "resistant" to the mechanisms of evolution, can be interpreted as networks with relatively large basins of attraction: the genetic operators endlessly cause new trajectories, but they all move towards the same attractor and therefore do not (or do not essentially) change the organism as a whole. Vice versa, biologically highly variable species must be interpreted as phenotypes with genomes that have no (or only few, small) basins of attraction. In this case the

[58] I abstract from the fact that, strictly speaking, only the genes themselves are directly changed, i.e., change their state in the network model, and that the functions are actually only varied through that, that is, indirectly.

matter of which system type is more advantageous depends on the environment - e.g., its variability.[59]

"Stabilities" of ecosystems can be interpreted accordingly, following a model by May (1976). If species are attractors in a "genetic space", then ecosystems are "supersystems" that developed from single species and that form trajectories themselves by the interactions of the species. Again, an ecosystem with many, large basins of attraction is highly resistant to disturbances, and vice versa. Analogous concepts regarding social systems should be obvious.

By pursuing these considerations a bit further it turns out, among other things, that some of the current criticism of classical Darwinism (which in fact has to be taken seriously) is no longer tenable. This criticism mainly refers to the - admittedly ambiguous - concept of adaptivity; it is summarized by Roth (1996, pp. 346; see also Gould/Lewontin 1979); I will go into that in more detail in order to clarify the fundamental importance that the systems theoretical concepts that are described here also (in fact especially) have for evolutionary theories:

(a) Many organisms have not essentially changed over the course of millions of years although their environment has changed; (b) on the other hand, there are organisms that did change in a constant environment; (c) many organisms survived just because they did just not adapt to their environment *in a highly specialized way*, while others died out because of their too specialized adaptation; (d) organisms of the same origin can develop differently in the same environment. From these results Roth draws the conclusion that the basic principles of classical Darwinism (now in form of neodarwinism) will have to be revised.

As a non-biologist I can not judge the *biological* content of the considerations of Roth, Gould, and Lewontin. From the point of view of a mathematical theory of adaptive and self-organizing systems, however, it can be assumed from what we have heard so far that these arguments are not as cogent as Roth and other critics of neodarwinism seem to think:

(a) If a species represents an attractor in the genetic space, and if this is in a basin of attraction that always takes different trajectories to the same attractor, then the constancy of a species even in a variable environment is not surprising at all, but is in fact a mathematical necessity. This results on the one hand from the convergence of the GA, which takes a species, provided that it has enough time for developing, to

---

[59] This quite obviously explains why there are species who remained unchanged for millions of years; in a constant environment such as, for example the deep sea, having great basins of attraction was advantageous, and so neither mutation nor crossover were able to change the attractors, and with that the organisms.

a specific state that is adequate with regard to its environment. At this point there are basically two possibilities, as will be shown in more detail in section 3.3 and chapter 4: (1) The genome of the species corresponds to a rule system that quickly takes the system - the phenotype - to a point attractor or to an attractor of such a small period that the phenotype does not essentially change anymore. Then the species stops changing for purely *mathematical reasons*, even if the environment is variable; our system's self-organization prevents further changes, even if it might be better for the species to adapt once more to the changed environment. (2) The genome represents a rule system of Wolfram class IV (or one that approximates that class). In this case, it usually takes a long time before attractors can be reached, *even if the genome does not essentially change because of the convergence of the GA.* The species then has to be variable for the same formal reasons, even if this is "actually" not required in a constant environment. Argument (b) is thereby revised as well.

(c) and (d): we have already seen in chapter 2 that adaptation can not be equated with the overly simple concept of assimilation. And we will see from the reflections on hybrid systems in section 3.3 and chapter 4 that specialized adaptations do not always have to be favorable at all, but that successful adaptation always means that the systems realize those rules (genomes) and states (phenotypes) on the basis of their self-organization that are adequate with regard to the environmental requirements because of their actual "history". On the one hand, very different solutions are possible and can be expected for purely mathematical reasons - argument (d); that is because the systems are, mathematically speaking, underdetermined. On the other hand, systems that do not adapt too narrowly but have a certain range of variation are often the most successful (argument (c)): in terms of AI research these are systems that largely operate with so-called "default rules", i.e., rules that are not very specialized, but that can be applied to a relatively large number of situations for exactly that reason. The systems thereby become relatively independent, at least from environments that do not change too radically; compared with that, highly specialized systems can be more successful in specific environments, but even slight changes of the environment can get the systems into trouble because then their specific rules do not apply anymore.

What do these considerations tell us? It may be true that Darwin's theory of evolution is in need of revision for biological reasons; this I cannot judge. From the point of view of systems theory, however, it turns out that Darwin's basic assumptions can in fact be justified by extending

his concepts with the insights of the theories of self-organizing and adaptive systems.

Thus, such an analysis of biological evolution uses the idea of the hybridization of systems, which has to be understood here as the logical layering of two rule systems, in contrast to the usual biological meaning of the concept of hybrids. With that, it becomes very clear how closely connected and intertwined self-organization and adaptation really are. The matter of self-organization is usually left out of consideration in classical Darwinism; as a non-biologist, I tend to agree with Kauffman's arguments against neodarwinists like, e.g., Dawkins (1987) on this point.

The numerous experiments that have been carried out with GAs in the meantime have shown in particular that the decisive factor for the convergence of the GA is not mutation but crossover. From a classical evolutionary biological point of view this is by no means a matter of course, since there the decisive role in evolution has been mainly assigned to the principle of mutation. Our experiences with GAs show that the part of selection-controlled exchange of single or several genes may have been underestimated, though mutations are of course very important too. In GAs they become especially important when the system is stuck in an "evolutionary blind alley": this happens when the elements (that is, the strings of symbols) have become largely identical and no new possibilities of combination can arise from crossover (in our own experiments this was not infrequently the case). In the long run, only mutations can lead out of such a blind alley; when they occur and the elements have again become significantly different, the much more effective crossover can again take over the actual optimization process of the elements.

From a sociological point of view it seems reasonable to interpret crossover as the cultural exchange of different societies, and the already mentioned predominance of crossover as a sign of the paramount importance of cultural contacts for sociocultural evolution. This is also expressed - from a psychological point of view - by Bateson when he points out that new ideas have their source mainly in the new distribution and recombination of ideas we got already (1979). In the following chapters we are going to find out to what extent these interpretations can be justified.

In order not to leave our analysis too general we are now going to look at a concrete example for the possibilities of experimenting with GAs. It is a program that was constructed by Ralph Kier and myself for the analysis of the iterated prisoner's dilemma, following a model developed by Axelrod (1987).

The name of the program, "prisoner's dilemma", comes derives from the fact that it is usually illustrated by the situation of two prisoners A

and B who are accused of a common crime. The prisoners are locked up apart from each other, and the police suggests a deal to each of them: if A confesses and B does not, then A is released (as a person who turns State's evidence, so to speak), while B can expect a severe sentence, and vice versa. If nobody confesses, the prisoners' actual crime can not be proved, and they are both given a minor sentence. If both confess, they are both given a heavy sentence, but not as heavy as would be given to the one who does not confess in the first case.

Since both prisoners can assume that this situation will never occur again, it is quite clear which strategy is objectively the most favorable for each of them: he must confess. If A does not confess, that is, if he behaves "cooperatively" to his accomplice, he runs the risk of being given the maximum penalty while the other gets off scot-free. If B also confesses, then A is given a quite severe sentence too, but at least a lesser one than in the first case. Thus, this "simple" dilemma can be easily solved.

Things look different when the situation is repeated - "iterated". In this case, A and B know how the other has behaved before and can adapt their own strategy accordingly. Since there are rewards and punishments for every action in each situation, the decisive factor now is not winning or losing in a single dilemma situation, but the overall result after several actions. That is why the "iterated" prisoner's dilemma is in fact rather complicated and cannot be solved by as trivial mean as in the simple case, which has made it the subject of much analysis.

Usually this problem is analyzed within the framework of the mathematical theory of games in a standard form that goes back to Axelrod and to the "inventor" of the prisoner's dilemma, Rapoport (Rapoport/Chammah 1965). Particular actions' results are noted - in relation to the opponent's actions - as numerical values; if we describe the betrayal of B by A as D ("defect") and the non-confessing of A as C ("cooperate") we get the following "pay-off matrix":

A/B	C	D
C	3	0
D	5	1

This matrix should be read in the following way: if both players behave cooperatively they get 3 points each; if A cooperates and B does not, A gets no points and B gets 5 points and vice versa. If both do not cooperate they get one point each. We should keep in mind that the players do not have any knowledge of the other's strategy while they decide on their own strategy; furthermore, "cooperation" and "non-

cooperation = aggression" are only defined from the point of view of the players and their relation to each other.

In a variant of Axelrod's model (1987) we have used a GA to analyze the evolution of cooperative and aggressive game strategies. The difference between our model and that of Axelrod is of minor importance here, since the connections between the theory of games and systems theory will be analyzed in more detail in chapter 4. Like Axelrod, we started out from the assumption that the behavior of a strategic player mainly depends on his knowledge of the previous moves of his opponent; this is often described as "look back" strategy.[60] Therefore, we had to encode a vector for the GA that included all the possible reactions to the previous moves of the opponent. For reasons of capacity we determined that the last five moves of each opponent should be considered; if we encode a cooperative strategy with 1 and an aggressive strategy with 0, and if we lay down the first move - that is, 1 or 0 - then we get a binary vector of the following length:

$$1 + 2 + 4 + 8 + 16 + 32 = 63.$$

For n moves that can be looked back on we get a vector of the length $\sum_{i=0}^{n} 2^i$

This, of course, results (in accordance with the combinatory logic with which CAs and BNs are computed, which has become familiar in the intervening period) from the number of possibilities at each move that is looked back on.

The probably most famous strategy for the iterated prisoner's dilemma, namely the strategy of "tit for tat" (TFT) by Rapoport, is an interesting exception here. This very successful strategy only considers the last of each opponent's moves and always plays in the same way as the opponent did in his last move. We experimented both with a standard GA and with an elitistic GA using different combinations. Two of these combinations shall be described in more detail:

We formed 20 strategy vectors of a length of 63 by random construction. In the first combination, each of these played against a fixed strategy; here we describe the results of playing against the TFT strategy. After about 100 moves the 10 best vectors were exposed to the genetic operators - crossover and mutation - and the 20 newly developed vectors played against TFT again. This was repeated approx. 100 times which results in a total of 100 * 100 moves. So, what we have here is a system that contains a double layered interaction, as it were: on the one hand, the strategies play against TFT and get their respective value; on the other hand, as is usual with a GA, new "elements" arise through genetic operators and selection, namely strategy vectors. Thus, if we

---

[60] In chapter 4, we will get to know examples for look forward strategies as well, also with the use of GAs.

interpret TFT as the environment of our system, we have a system whose elements must prove worthwhile in the struggle with their environment, and which will gradually attempt to improve their efficiency with the mechanisms of the GA. Fig. 3.2.1. a and b show some typical results[61]:

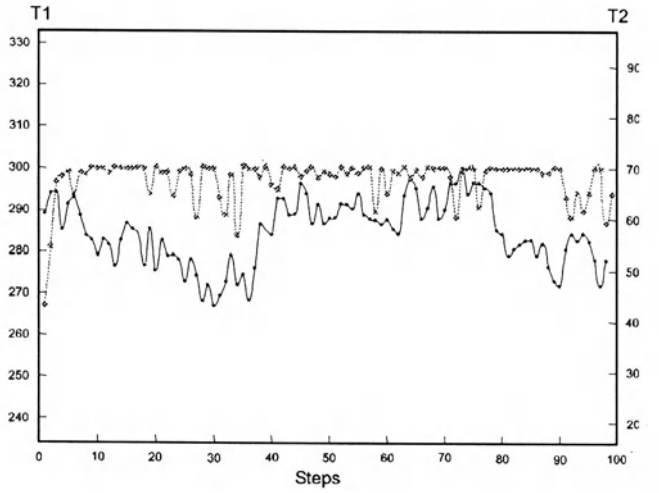

Figure 3.2.1.a

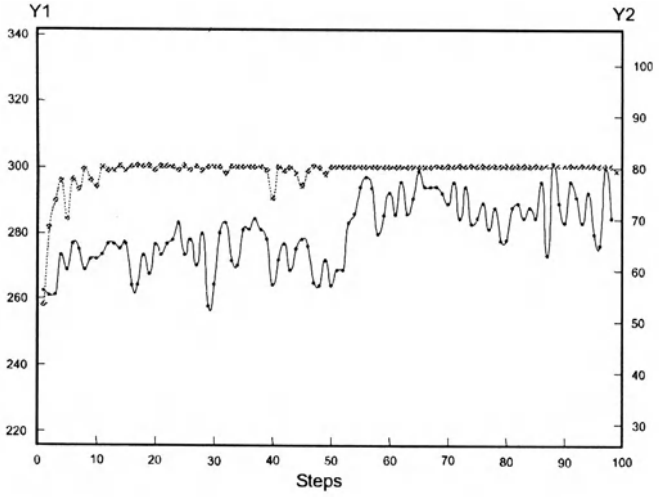

Figure 3.2.1.b

In both cases we used an elitistic GA, which means that the best strategy at any one time remained unchanged and was taken into the next

[61] Since our system is stochastic, relevant results can naturally only be obtained as mean values from several runs.

"tournament". To avoid any misinterpretation, it has to be noted that we determined the graphs of proceeding according to the mean values of the 10 best vectors and their "cooperativity" (proportion of the component 1 in the whole vector).

As we can see in Fig. 3.2.1.a, the strategies' mean values group around a score of 300; this means that relatively cooperative strategies establish and assert themselves quite quickly. In fact, a score of 300 is just the maximum value that a cooperative strategy can achieve against TFT. Compared to that, aggressive strategies would always come out worse, since TFT responds to aggression with own aggression, and can therefore attain only 1 point at a time, that is, an average of 100 points per tournament. So, selecting strategies by their success in the tournaments quickly results in an adjustment. However, we could see that this does not happen homogeneously, but that there are instead continual deviations, which make the jags in the graph of points. The reason for that is the relatively high mutation rate of 10%, which keeps on bringing up more aggressive strategies.

Obviously, the cooperativity values' graph of proceeding is not so smooth, although it approximates the other graph's fluctuations. Here it becomes noticeable that the immediate selection orients itself by the achieved score, and that the degree of the players' cooperativity comes to fruition only indirectly. In any case, we can see that strategies with a cooperativity proportion of at least 50% outweigh the others by far; this corresponds to the degree of cooperativity of TFT. So, in the end the results is a stabilization with a cooperative tendency as well.

Fig. 3.2.1.b clearly illustrates the influence of the mutation rate. The mutation rate has been reduced here to 0.025 and we immediately see that both graphs of proceeding show much less fluctuation. Both graphs of proceeding stabilize at 300 points and a degree of cooperativity of approx. 50, respectively, and deviate only insignificantly from this value. Thus, with due caution we can say that the pair (300,50) is something like an attractor. Naturally, the low mutation rate constantly draws the graphs of proceeding somewhat away from this point; however, the selection brings the development of states back to it with the same constancy.

In the second version, the strategies do not play against a fixed strategy that is preset "externally", but against each other; again the selection is carried out according to the highest score, and the best strategy remains unchanged and is taken over. As a consequence, the system does not have to deal with a constant external environment, but is literally its own environment, in fact a highly variable one: since new strategies are formed by the genetic operators after every tournament, it is also possible for other strategies to be successful. Axelrod (1987) aptly

describes this situation as that of a "moving target", namely the orientation by an environment that constantly changes precisely because of that orientation. This elucidates an insight that Axelrod had already formulated in 1984, namely that there can be no absolutely best strategy. A strategy is the best only in a specific situation. Fig. 3.2.2. shows a result:

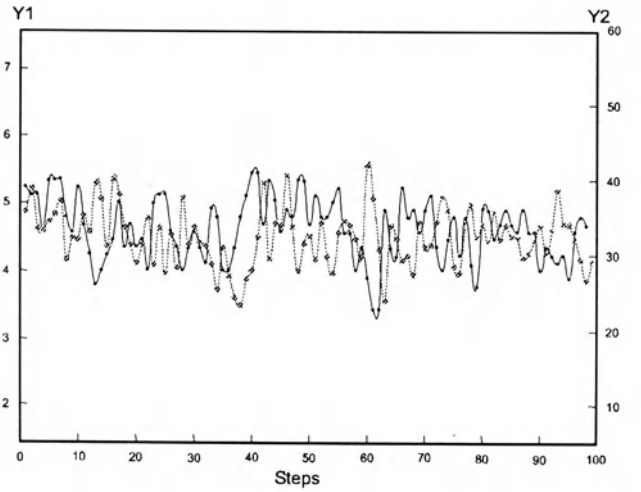

Figure 3.2.2.

Here the picture looks completely different. For a better understanding it has to be mentioned that in this instance significantly more points can be achieved on the whole, because every strategy plays against each of the other 19 strategies. Thus, with a maximum score of 5 per move a strategy can achieve a maximum of 500 * 19 = 9500 points in a tournament of 100 moves. Here the graph of points again ranges in the scope that could be carried out against a strategy like TFT, namely 300 * 19 = 5700. However, the fluctuations are much greater and tend to stay below the possible value of 5700. This is because in the case of 19 different strategies there are always some relatively aggressive ones getting a chance to achieve higher scores by "exploiting" cooperative strategies; in the medium term, however, the other strategies adapt themselves to that, and as a consequence they all get low scores. Accordingly, the cooperativity values are much lower here than in games against TFT.

It has to be mentioned that we again used a high mutation rate here. However, reducing the mutation rate does not have such a clearly stabilizing influence in this case - precisely because of the heterogeneity of the individual strategies.

We are going to deal with the connection between game theory and systems theory more systematically in the 4th chapter, and therefore I will not comment any further on the contents of these results here. The intention behind this example was simply to show that, and to what extent, a GA can induce a system dynamics by itself; we should keep that in mind - see the quotation of De Jong at the beginning of this subchapter - when a GA is combined with other systems.

Any system that is able to respond to its environment in any form - however weakly - is necessarily adaptive. In the strict sense of the word, however, this is only true of systems that develop by rule variation; we have tried to elucidate that with the example of biological organisms. The principle of rule variation can be realized in a particularly transparent and technically manageable way by use of hybrid systems; in the following we are going to take a special look at this type of formal system.

3.3. HYBRID SYSTEMS: DYNAMICS AND METADYNAMICS

> "To parallel this (i.e. learning, J.K.) in a computer system, a higher-level program would have to have some way of creating changes inside the 'demon' that is carrying its program out. This is utterly foreign to the present style in computer science of implementing one level above another in a strictly vertical, sealed-off fashion. The ability of a higher level to loop back and effect lower levels - its own underpinnings - is a kind of magic trick which we feel is very close to the core of consciousness. It will perhaps one day prove to be a key element in the push toward ever-greater flexibility in computer design, and of course in the approach toward artificial intelligence. In particular, a satisfactory answer to the question of what 'understanding' really means will undoubtedly require a much sharper delineation of the ways in which different levels in a symbol-manipulating system can depend on and effect one another. All in all, these concepts have proven elusive, and a clear understanding of them is probably a good ways off yet." (Hofstadter/Dennett 1981, 380)

When Hofstadter and Dennett expressed these programmatic insights, the field of research was still widely dominated by so-called "symbolic AI," - that is, by the attempt to grasp intelligence with rule-based systems that do not change after their construction and, above all, that operate on only one logical level. In the meantime, what Hofstadter and Dennett called "a kind of magic trick" has almost become taken for granted in the branch of research that deals with hybrid systems, namely programs that operate on different logical levels and "loop back" to refer to lower levels from higher levels. So there has been a kind of paradigm shift, not least because of the Artificial Life context. Like Hofstadter and Dennett, I am convinced that here we have at least *one* key to the analysis of

phenomena as complex as the mind or society; however, the following passages will also show that the ideas connected with this problem are indeed very "elusive", and that we are still at the beginning of a probably long journey. Therefore, only a few completed results can be presented in this subchapter; at first, we have to conceive the path in general. In any case, the first moves on this path have been taken, and at its end there may be the "clear understanding" that Hofstadter and Dennett were hoping for.

3.3.1. Rule Variation, Environment, and Hybrid Structures

The systems that the cultural and cognitive sciences are dealing with have been characterized according to various aspects in sections 1.3 and 2. In the following sections we are going to look at them in more detail; however, it should have become clear already that all the properties in question - adaptation, self-modeling, learning, self-enlargement, etc. - can only be accomplished if the rules of interaction between the elements of a system are variable. We could see in section 3.1. that pure self-organization can lead to highly complex system dynamics; however, those dynamics are blind to their environment, as it were, and so they cannot explain how systems can respond to their environment in a goal-directed way - especially if the environment itself is variable.

Bateson's analysis of the concept of learning clearly pointed to the fact that adaptive systems that are capable of varying their rules of interaction must actually have more than one set of rules. Bateson, who followed the theory of logical types by Russell and Whitehead, understood this as a hierarchical nesting of different rule systems (in which levels III and IV are, of course, not precisely worked out). If we abstract from these problems, we can characterize an adaptive system by the fact that it must have at least two different, interdependent sets of rules. There are of course various possibilities: one set of rules might modify the other but is not modified itself; both may mutually influence and vary each other; and finally, both can maintain their logical form but change particular parameters of the other set of rules.

As we could see, even those artificial neural nets that learn have rules for the interaction between their artificial neurons and their rules - their learning rules - for the modification of their interaction parameters; modifications of the linear standard rules for the programs' neural interactions have also been tested - e.g. by introducing multiplicative rules (Pearson et al. 1995). The realization of this is especially effective when the different sets of rules are implemented into different parts of

the program and these parts represent different program types. When this is the case, we speak of *hybrid systems*.

Both the term and the concept of hybrid systems have, to my knowledge, originated within the context of AI research, especially under the influence of two competing tendencies, namely the so-called symbol processing approach on the one hand (usually carried out by rule- or knowledge-based systems) and, on the other, the connectionistic approach, which mainly works with artificial neural nets. It suggested itself to see these approaches not as a logical alternative, but instead to integrate the two approaches. From this, the first considerations on hybrid systems originated (e.g. Helm 1991). However, these considerations did not get very far, since it was not really clear which part of the human - and then artificial - intelligence was operating by processing symbols, and which was more connectionistic. In contrast to "pure" program types like NNs, CAs, BNs, and GAs, hybrid systems were hardly analyzed systematically. Only recently has a larger number of scientists begun to study the possibilities of hybrid systems (Goonatilake/Khebbal 1995), and it is certainly no accident that the coupling of neural nets with other programs is at the center of these studies (Belew et al. 1992; Pearson et al. 1995). The techniques of "genetic programming" developed by Koza (1992) and others should also be mentioned in this context; their purpose is usually to use GAs for the modification and extension of LISP programs.

Before we go into the techniques and problems of hybrid systems in more detail, I would like to illustrate the relevance of that system type for the cultural and cognitive sciences, that is, for systems that can change themselves, once again with a famous scientific example. It is the Chomskyan model of a universal grammar (e.g. Chomsky 1965) and the concept of (children's) language acquisition developed therefrom.
Aoun (1994, 122) describes the Chomskyan program in a concise, systems theoretical way:

> "The fact is that there is something about speakers of English, for instance, which differentiates them from speakers of Chinese. That is some state of their minds. It is a steady state that they have attained sometime in childhood. They began at some initial state and went through a series of states until they reached that state."

That sounds very much like a self-organizing system that reaches an attractor state from an initial state through its own dynamics. But if we look at it more closely, we can see that we are dealing with very complex interconnections between self-organization and adaptation here:

Following Chomsky and his school, language acquisition is realized in two dimensions. First, the universal grammar determines general frame conditions that define, so to speak, that the developing language system

really is language; these frame conditions are the linguistic universals. Second, the speech community in which the child grows up determines which specific rules have to be built up for a particular language to develop. These rules are both ontogenetically and phylogenetically variable, which means that children - and adults too - can vary the rules of their language in the course of their "speaker life", while their specific language can change considerably change over the course of generations as well.

Pinker (1994, 270) illustrates this point with the following example: it seems that most linguistic universals have the form of implications; thus, they are themselves rules that determine the structure of any specific language. For example, if a language has an NVO word order - noun, verb, object - like German or English, then the interrogative particles in an interrogative clause are always at the beginning of the sentence, and the language has prepositions. If, in comparison, a language has an NOV word order, then it has postpositions and the interrogatives are usually at the end of a sentence. Thus, this linguistic universal does not determine *that and if* languages have NVO or NOV word orders, but it lays down consequences for each of the possibilities. To quote Goethe's Faust: "*Beim ersten seid ihr frei, beim zweiten seid ihr Knechte*" (The first is free, the second's slaves are we).

If we translate this general conception into a formal model, it suggests itself to think of a BN, in fact - referring to language acquisition – a BN that has to build itself up yet. The network does not "know" any more than that it "is" a network; furthermore, there are some *a priori* valid rules of combination that are implemented into the network from the beginning. With that, the net begins its construction and tests some rules that it develops for itself. Here the introduction of an NOV word order rule is possible with the categorizations N, O, and V belonging to the universals - during its construction the net takes into account that it has to introduce different elements or elements in different states. If the network's "linguistic environment" has realized the NVO rule, the net has to "learn", that is, it varies the NOV rule in favor of the NVO rule. It follows automatically that now the above mentioned rules about prepositions etc. have to be realized, since this has been implemented *a priori*. So now propositions can be introduced as further elements, and so forth, until the actual development is completed; the net has developed its rules - and elements - sufficiently to meet the requirements of the linguistic environment adequately; and the net can reach an attractor state by simply applying these rules - Aoun's steady state. It has to be stressed

once again, however, that this development only takes place because of the continual interplay of self-organization and adaptation.[62]

With this little illustration I certainly did not intend to claim that the structure of a linguistic system has to be interpreted as a hybrid system in the above sense; that would require the assumption of something like two modules in the brain, one of which would have to include the rules of the specific language together with strictly determined constants of the universal grammar (e.g., the fact that all languages have interrogatives), and the other would have to have rules for the variation of the specific language's rules as well as linguistic universals that are themselves formulated as rules, as in the above example. I think that such an assumption is rather improbable (cf. Müller 1991), and could besides hardly be derived from programming experience. What I wanted to illustrate is that we have to assume at least two logically distinct sets of rules here, one of which influences the other. I gladly leave the question of *how* this is realized in the brain to neuro- and psycholinguists.[63]

Naturally, the developmental process of our network is not completely finished, and the attainment of the attractor state does not mean the end of the system's history. Adults feel obliged to vary their linguistic rules, too, e.g., in new social environments and/or in the face of intercultural contacts. Adaptation presumably outweighs self-organization in the actual developmental process, since, to my knowledge, research shows that the variability of the single languages is probably much higher than the *a priori* determinations of the universal grammar - there can be more rules varied than is determined by the universal grammar. At the end of the child's developmental process self-organization predominates, since the linguistic system can basically realize its different states by the rules it has built up to this point without changing them considerably. These subtleties, however, may be left to linguists with an interest in systems theory. What I wanted to illustrate here was "simply" that, and to what extent, every important area of the cultural and cognitive sciences that is to be described in terms of systems theory necessarily requires a modeling that is hybrid in a *logical* sense, i.e., one that introduces and differentiates more than one set (you could say instead: "multiple sets") of rules.

---

[62] I abstract here from many details - like Chomsky, Aoun speaks, e.g., of the "state of mind", while I have confined myself to the subject of language as a system here. With that I rather follow Schnelle's considerations (1991). Furthermore, Chomsky originally understood something different by linguistic universals (see below, section 5.2); I adopt the categorization of his follower Pinker (loc. cit.) here.

[63] Pinker (loc. cit.) seems to assume such a modular structure of the brain with regard to language - and not only to that.

It has to be mentioned in this context that there is yet another concept with which the phenomena that we are examining here can be logically grasped: in a similar context Farmer (1990) speaks of "metadynamic" systems (see also Dupuy/Varela 1992; Bourgine/Varela 1991a); what he means are systems that can not only change their states, but also their specific system parameters and even their architecture; quite obviously, Farmer is thinking of the logical relations between the system elements, and partly of the rules of interaction as well. He illustrates the concept of metadynamics with models of the immune system, artificial neural nets, autocatalytic systems, and the classifier system developed by Holland (Holland et al. 1986). However, Farmer does not speak of logical levels here, but of differences in the time scale: a metadynamic system changes its states in the short term, some of its parameters in the medium term, and its architecture only in the long term. This unquestionably makes sense for natural systems; in the case of social systems it is not necessarily so. Farmer's rather general considerations do not contrast with the analysis presented here, though: metadynamic systems are always hybrid in a logical sense, as could be seen from the example of artificial neural nets. That is why I will use the terms hybrid and metadynamic system synonymously in the following, without further consideration of the time scale mentioned by Farmer.

Let us now take a closer look at hybrid systems in the narrow sense, that is, with regard to the technique of programming. A hybrid system in that sense consists of two or more individual programs that are coupled with each other. As I already mentioned, there are very different possible ways of doing this; basically, we can distinguish two main types:

On the one hand, it is possible to couple the two subprograms "horizontally". This means that different tasks are assigned to each subprogram, which are handled separately and in a specialized way - on the principle of the division of labor, as it were. Neither of the programs is sub- or superordinate to the other; they either run their respective tasks in continual succession, or the maker or user of the program decides at what time he wants one or both of the different tasks to be carried out.

The system THEOPRO, which was described in Klüver 1995, is an example of such a horizontally coupled hybrid system. In this system a rule-based system is coupled with an (interactive) neural net; among other tasks, the rule-based system is able to build up a matrix that is subsequently the NN's basis for several simulations. The results of the simulations can be returned to the knowledge base of the rule-based system so that this knowledge base can be regularly enlarged.

In a vertically coupled hybrid system, one subprogram is modified by the other, i.e., it is influenced "from above". Usually this is done by

improving the one program by means of an optimization program that is not itself changed. A relatively simple example of this is the CA that is described in more detail below; it is modified by a GA, e.g., by varying specific parameters, namely probabilities of state transitions.[64]

These two rather simple basic patterns can be varied and combined in various ways. In the meantime, the mentioned program THEOPRO is additionally supplied with a GA that optimizes the interactive net - more precisely: its weight matrix (Stoica 1997). Thereby, a learning mechanism has been added to the whole program, which makes it a horizontally-vertically coupled hybrid system. Accordingly, we coupled a CA with an interactive net and a GA that alternately optimizes the network and the CA; in the following chapters I am going to describe this system in more detail. Obviously, it is a horizontally-vertically coupled hybrid system as well.

Vertically and horizontally-vertically coupled hybrid systems are particularly interesting for the modeling of systems that we are dealing with here. The common logical characterization of the adaptive, learning, etc. systems that we have been looking at so far is that the different sets of rules from which they generate their dynamics always operate on different logical levels. This was illustrated especially in connection with Bateson. That is why I will now concentrate on vertically coupled systems; in the following, hybrid systems will always be understood as vertically coupled systems, except where explicitly noted otherwise. According to the usage of logical semantics (Tarski 1956), I will call the subprogram that modifies the other the *meta-program* and the modified program the *base program*. However, we shall see that under certain circumstances the base program can also modify the meta-program, and that things are unfortunately not as simple as with the strictly separate logical levels of Russell and Tarski; but for the moment I will abstract from this point.

Now what is meant by "rule variation"? Except for the little digression on artificial neural nets, I have used this concept without defining it yet, and left it to the intuition of ordinary language. But if we look at it more closely, which is inevitable for the construction of programs, we can distinguish various different possibilities of rule variations, and the following list is probably not even complete:

(a) Presumably the logically simplest case is to either apply or annul a rule.

---

[64] In a similar context Goonatilake and Khebbal (1995a) speak of "intercommunicating hybrids" and "function-replacing hybrids"; these concepts do not completely agree with the concepts of "horizontal" and "vertical" and they only include rather special cases.

(b) A modification of (a) involves having the application of a rule controlled by a probability parameter practically, this means that the rule is in force, but takes effect only according to a specific probability. It is plausible (and has even been described in section 3.1.1) that a variation of this probability parameter could really change the dynamics of, e.g., a CA. Another example is the variation of weights in artificial neural nets that was sketched above; in the noted coupling of a CA with an NN, we have explicitly made use of the similarity of the variations of numeric parameters (see below).
Of course, we should also mention the case that was analyzed in connection with the FP parameter (3.1.), in which a stochastic system has several options per rule. A variation of such rules can either mean that the probabilities of the single options are modified, or that the number of options itself is changed. As we could see, rule variation often entails a fundamental change of the system dynamics in this sort of case.

(c) Logically, the rules consist of an antecedent part - the conditions - and a following part. As a consequence, another case of rule variation lies in varying either the antecedent part, the following part, or both parts of the rule. In a CA, e.g., the transition from a Neumann neighborhood to a Moore neighborhood would be an antecedent variation; in the case of totalistic rules for CAs, a variation of the mean values of the neighboring cells could be classified accordingly.

(d) If it is possible, as in (a), to factually reduce the number of rules by annulling a rule, then there is of course also the opposite way of increasing the number of rules, which means that the program "itself" introduces and tests new rules. We have realized that for CAs and interactive neural nets (Stoica 1997); the sociological aspects of this "self-enlargement" will be dealt with in more detail in section 4.5.

(f) Finally, all these separate cases can be combined. We have experimented with such combinations too, and it is naturally quite difficult to keep track of things there: which modification caused which effect on the behavior of the system? These hybrid systems are not very suitable for "clean" experiments in the classical style of the experimental natural sciences, since there are always several variables changing at the same time. At any rate, I am afraid that in social reality it is always, or at least often, the case that changes simply do not occur as controllable "one variable variations"; this shows itself, e.g., in the development of large-scale technologies, as can be seen from their frequent "malfunctions". Thus, experiments with such complicated hybrid systems instead have the function of obtaining a global impression of what can happen when several variables are varied at the same time - that is, of *how* complicated social reality really is.

I would like to present an example of a relatively simple hybrid system now in order to clarify further what we have said so far. This system is the hybrid CA that was already mentioned in section 2.2. (see also Klüver 1996):

As I said before, we intended to simulate a society that develops from a homogeneous rural population to a socially differentiated class society. The single classes are symbolized by different states of their cells; the neighborhood rules are totalistic, that is, only mean values of the neighborhood cells' states are formulated as conditions. The neighborhood itself is a Moore neighborhood. The rules are stochastic, so that the transitions from one state to another (i.e., upward or downward mobility from one class into another) can only be given in probability parameters. Sociologically there is no alternative, because even a concept like social permeability, which is modeled by the transition rules, only contains probabilities for the social fate of an individual.

This CA, which was given the pretty name SOZION, is coupled with a GA whose task is on the one hand to vary its rules with regard to the *probability parameters*, and on the other to turn single transition rules on or off. Thus, according to the above systematics, SOZION is a combination of (a) and (b). Since the purpose of SOZION was to check the different consequences of different social mobility rates for "favorable" final results, and mobility in the formal sense can mean ascent as well as deterioration or standstill, we needed a combined, though still simple, rule variation.

The "units" of the GA are ternary coded vectors. Here we have to distinguish between two rule variations: for the variation of the probability parameters, 0 means that the probability rate of a specific rule remains unchanged; 1 means that the rate is increased by a specific value (e.g., 0.05); and 3 means that it is decreased by the same value. For the variation of the antecedens parts, 1 means that the actual rule is in force - with the respective probability of course; 3 means that there must be an upward mobility; and 0 means that nothing happens (case (a)). Thus, every component of the GA vector is assigned to one and only one probability rate of a rule, or one and only one antecedens part of a rule. We usually run SOZION with a standard GA, i.e., not with an elitistic one, and with both genetic operators, that is, crossover and mutation.

A GA of course makes no sense if it does not contain an evaluation function for the selection of the "most favorable" solutions. As a consequence, we first have to define which solutions are favorable for a social system. For this, we make a rather simple assumption whose basic idea is borrowed from the functionalist tradition, especially from functionalist systems theory (see above, section 2.7.): social systems have to process specific functions to fulfill various environmental

requirements; among these functions are, e.g., the production of food and other material products; joint action towards other social systems; ensuring of social integration and social harmony; and so forth. Evidently, in accordance with Luhmann's conception, we have also counted the individual's mind as among the environment of social systems. This list of functions is neither very systematic, nor does it claim to be exhaustive; primarily, our model construction aimed to show the development of social differentiation as the optimization performance of a social system with regard to specific environmental requirements. Thus, according to the considerations we presented so far, our artificial system should be both adaptive and self-organizing.

We furthermore assumed that the social - stratificatory - differentiation has to be interpreted as the development of the different social classes' taking over of single functions on the principle of the division of labor. Thus, ignoring technological developments, we determined for every class i and every system function j a "contribution value" B_{ij} that determines how much this class i contributes to the function j. It is evident that the number of class members is also very important for the efficiency of a class for the system as a whole, and so we defined the number of members of the class i as A_i. Finally, not all functions are of the same importance to the system; therefore, we determined F_j as the "value of the function" for every function j. As a result, we obtained the following computational formula for the "system value" W:

$$W = (\Sigma A_i * B_{i1} * F_1, \Sigma A_i * B_{i2} * F_2, \ldots, \Sigma A_i * B_{in} * F_n) \qquad (3.3.1.1)$$

So, W is an n-dimensional vector whose components are the sum of the members of all classes that contribute to one of the several functions, each multiplied by the actual contribution values and function values. As a result, the dimension of W is entirely determined by the number of functions (and not, e.g., by the number of classes). Since the functions need to be fulfilled to come up to particular environmental requirements, the environment is also described as an n-dimensional vector whose components represent, so to speak, the maximum level of each specific function that the system has to perform. If we call this environment vector U, then

$$d = |U - W|$$

is the distance between the performance of a system and the environmental requirements, and therefore is a degree of the system's - environment related - capability to respond more or less adequately to the

environmental requirements. The computation of d follows from its Euclidean distance.[65]

We have thereby completed the definition of the evaluation function. Of course, the criterion for the selections of the GA is the minimization of d, that is, a maximum approximation to the environmental requirements. Now here is a description of how SOZION works:

The system starts out with a homogeneous rural population. The rule system of the CA can either be generated at random or according to specific criteria. Thereafter, another 20 ternary coded rule vectors are generated at random. These vectors are now "applied" to the CA's set of rules, which means that, depending on each coded element, the rates of probability are either kept constant, increased, or decreased, and the rules are either in force or they are not. As a result, we get a total of 20 new sets of rules, i.e., 20 CAs. These are tested now, that is, they execute a specific number of steps - usually about 10 - and then the value W is computed for each of them. After this, the best sets of rules are selected, and the GA takes up its operation again after the known principle.

There are several things to be noted about this program: first, it should be obvious why I introduced the concepts of base and meta-program above. The CA is in this case the base of the whole program, because only the runs of the CA can be seen as simulations of the social differentiation, and only the states of the CA or CAs decide on what is a favorable solution and what is not. In itself the operations of the GA are meaningless; they get their meaning within the whole program only by the fact that they are applied to the sets of rules of the respective best CA. Correspondingly, the genetic operators only affect the ternary vectors. Therefore, the terminology used above suggests itself not merely for logical reasons.

Second, it is plain to see that such a program is rather unrealistic. Apart from the fact that the evaluation function is certainly much too simple, it is apparent that social systems do not permanently change their rules, especially not by increasing those rules twentyfold. With this in mind, the question of *what* is actually simulated by such a hybrid system

---

[65] The Euclidean distance between two vectors V and W is computed in the following way: Let $V = (v_1, \ldots, v_n)$ and $W = (w_1, \ldots, w_n)$. Then $d = \sqrt{\Sigma_i (v_i - w_i)^2}$. The so-called Hamming distance, which is often used in computer sciences, is not so suitable for our purposes, since it is only applicable to binary coded vectors.

arises; however impressive the system's optimization performances may be, they have got nothing to do with any kind of reality.[66]

However, we should not let this first impression deceive us. Strictly speaking, there is always just one base system and not twenty. The hybrid system's operations have to be interpreted as the continual changing of one "real" CA by continually orientating to 20 "virtual" CAs; these 20 virtual CAs are possible developments of the real CA, and only one of them is selected; this process subsequently yields the initial states for the further real or virtual developments. To elucidate all this in sociological terms: the virtual CAs are generated by a hypothetical adoption of interculturally gained stimuli (crossover), together with self-innovations (mutations): of these new possibilities they only select the one that is the most favorable with regard to their own goals (or that seems to be the most favorable, since this cannot be predicted for the long term).

Both the *continual* variation of the rules and the construction of 20 virtual CAs only serve to speed up the whole process in this program. In the following we are going to deal with programs that behave much more realistically in many respects, in that they in principle neither permanently change their rules nor vary them all at once. The purpose of SOZION was first of all to test and elucidate the fundamental possibilities of hybrid programs for the modeling of self-changing social systems. And for this purpose SOZION is very well suited.

Third, with a system like SOZION it can be clearly shown why adaptation must not be confused with the overly simple concept of assimilation. Of course, SOZION orientates its developments to its environmental requirements, and thus to external criteria. However, this does not predetermine a single solution. One can rather see from the results that very different "societies" can yield equally, or at least comparatively, good solutions. It would obviously make no sense to speak of these different systems as being "adapted to the environment" since strictly speaking they adapted to d, and they did that by use of algorithms that are not particularly simple. What we have here is therefore a very illustrative case of adaptation: a self-organizing system responds to environmental requirements by a rule variation whose meaning for the system is checked by its selforganized dynamics; if

[66] SOZION usually starts with d = 0.6 or something like that. After the GA has run about ten times, it almost always reaches d = 0.08, which is another proof for the GA's efficiency. This often is a final point for SOZION, because the vectors have by then become too similar and the crossover does not lead any further. This value can even be improved when using relatively high mutation rates, which is quite arduous though. In chapter four I am going to describe a system that is even more efficient; this is achieved - see below - by the variation of "meta-parameters".

necessary, another rule variation follows, etc. However, the stress is on the intrasystemic dynamics, because the changes of the CA's states by application of its respective rules, as well as the rule variation by application of the genetic operators to the GA vectors, are processes that take place on the basis of the logic of the whole system - the hybrid system.

The reason for the fact that there are any different solutions with comparatively good values at all lies of course in the degree of freedom according to which the system can vary. In the basic model of SOZION we have treated only A_i and F_j as free variables, while keeping B_{ij} constant. This is also not very realistic, because the values of the contributions of a class to the system value can change too. In a society like Germany, e.g., the production of food is probably not as central as in mainly agrarian societies, which entails a reduction of the corresponding F value; the farmers' growing dependence on agricultural chemistry and machinery reduces the farmers' B value accordingly. Experiments in which these values were also treated as free variables, i.e., as set by the system itself, showed of course an even greater heterogeneity of favorable solutions (see below, chapter 4.). In connection with the critical examination of some arguments against neodarwinism in section 3.2, we have already seen that this is practically standard in the case of complex adaptive systems.

Fourthly and finally, it seems to be much more complicated to determine the dynamics of a system like SOZION in comparison with the dynamics of the systems we looked at in sections 3.1. and 3.2. Strictly speaking, we are dealing with a permanent overlap of two system dynamics, namely of the base and the meta-program. There can be no general decision on whether or which one of the two is predominant.

I would like to elucidate this problem with the example of a variant of SOZION, namely the hybrid program SOZGEN, which was used by Jörn Schmidt and me for the analysis of social processes in a seminar in the winter semester of 1995/96.

SOZGEN - "the genesis of societies" - had been constructed as a variant on SOZION to clarify in a model the claims of Berger and Luckmann (1966) - which also comprise a model - about the genesis of social rules. SOZGEN usually starts out with two individuals of opposite sexes who form a purely agrarian mini-society together. Like in SOZION, there are environmental requirements, and a GA that operates accordingly. In contrast to SOZION, there are "reproduction rules" and "mortality rules" in SOZGEN: two individuals of opposite sexes produce new individuals on the basis of a particular reproduction rate, and the individuals have a specific maximum life expectancy; the older they get, the greater is the probability of their death, and after a specific time they

die. All other differentiation rules are the same as in SOZION; in addition, we should mention that normally the reproduction and mortality rules cannot be modified by the GA. It would be no problem to introduce this possibility too, in principle, but that would not aid the clarity of the model, which is already quite complicated as it is.

Additionally, SOZGEN is able to switch the GA on or off. This is done in the following way: the program computes d as usual, and if d is small enough it switches the GA off; if d is relatively large, the GA is switched on again. Of course, it has to be determined beforehand what is meant by "small enough" (see below, section 3.3.2.).

It is possible to start SOZGEN with a set of rules that corresponds to a low Wolfram class, e.g. class I, which would lead to an attractor state relatively quickly (the GA is switched off at the beginning in order to give our original couple the chance to try out the social rules first). Accordingly, the dynamics of the CA predominates in this initial phase. If the attractor state shows itself to be favorable, the system stays with it, since the GA does not have to be switched on. But this is not usually the case, because the system has to take the dynamics of the population into account, and one has to bear in mind that different system functions cannot be fulfilled adequately by all individuals at the same time. That is why in most cases there will be a rule variation by the GA after a short period of time, which usually causes an increase in social mobility. If the variation is only a slight one, the system can approach the same attractor again - it remains with its trajectory in the attractor's basin of attraction - or else it moves towards another attractor. However, this happens only if d could be minimized sufficiently on the basis of the first rule variation. For the most part this is not the case, and then new and more radical rule variations can take the system to completely different trajectories. Of course, these changes do not happen by accident, but instead according to the convergent logic of the GA, so that the latter introduces the predominant dynamics here.

The purpose of this subchapter was simply to provide an outline of the concept of hybrid systems. In section 3.3.3. this concept will be systematically analyzed, so for the moment we can turn to another problem, namely that of the "small enough d".

3.3.2. Digression: Conditions of Rule Variation or: Quis custodiet ipsos custodes?

As I have said before, the GA in SOZGEN can be switched on or off depending on d being "too large" or "small enough". The program's criteria of whether and when the rules of the CA are modified by the GA are of course externally predetermined by the programmer. Compared

with this, there is of course no such external determining of criteria in social systems; here the system decides on whether, when, and how radical the social rules are (or shall be) changed. Since this certainly does not happen arbitrarily, we actually need yet another set of rules, one, in fact, that controls the meta-rules of rule variation. These meta-meta-rules can be historically variable too, and so we need further rules to control these, that is, meta-meta-meta-rules, and so on. As a consequence, we get into the classical problem of an infinite regress; it is trivial that this neither occurs in real social systems, nor can it be implemented into a necessarily finite program. And so the question arises of how this problem can be solved in a formal model without explicit external predetermination, which is necessarily ad hoc. I would like to point to the fact that that there are "3rd level rules" already in SOZGEN, namely the 3rd level rule to switch the GA on or off, as well as an instruction according to which SOZGEN changes the mutation rate of the GA depending on the speed of the optimization processes. Thus, the program controls the tempo of its variations itself in many respects, namely by switching the GA on or off and by varying the mutation rate.

The problem of an infinite regress is not new and has been frequently dealt with from various angles. In logical semantics the problem arises - similar to our case - because of the logical layering of different languages that refer to each other; it can only be solved *practically*, by the postulate of natural language being the "highest meta-language" (cf. Apel 1979) which has to be characterized as a logically incomplete system. It is more obvious for our sociological purpose to refer to the dialectical figures of thought with which Marx tried to solve a similar problem as regards content. Naturally, Marx did not speak of rule variation, but of the alignment of the *Produktionsverhältnisse* (relations of production) with the level of the *Produktivkräfte* (forces of production); the determining of the *Überbau* (superstructure) by the *Basis* (base); and of the historical mechanisms of revolutionary class struggles. However, if we abstract from this specific terminology and the narrowed view on social dynamics that is connected with it, we may characterize his solution in the following, more generalized way:

On the one hand, there are historical mechanisms which are constant, namely the necessity of aligning the relations of production with the continually developing forces of production. Thus, we have the forces of production as independent, and the relations of production as dependent variables. Together the actual levels of each reveal a global state of the social system. Since the forces of production are continually developing - their postrevolutionary "*Entfesselung*" (unleashing) - this is not an attractor of the period 0. Instead, the system continually alternates between Wolfram class IV and III. The development of the forces of

production prior to a revolution generates new states which aggravate the whole situation without varying the rules - d increases. This goes on until a critical value is reached and the system switches the GA of revolution on, so to speak, *on the basis of its realized states*, which opens the possibility of rule variations - in this case violent ones.

In order to explain this mechanism, which he did not model very precisely, Marx falls back upon the role of class consciousness, which on the one hand is determined by social circumstances, and on the other hand triggers and intensifies class struggles by its revolutionary change (MEW 37/463). Thus, the system produces states through its own dynamics that have a changing effect on the mind, while the mind in turn affects the changing of states and, finally, the variation of the rules. Since Marx, in contrast to, e.g., Luhmann, unconditionally regarded the problem of a (class specific) mould of the mind as the genuine business of social theory, he was able to develop a model in which the social system itself, by the changes of its states, produces the criteria of when and how rule variations shall and must be carried out - formally, an extremely elegant solution.

It is unquestionably possible to adopt Luhmann's distinction of psychic and social systems and at the same time follow the Marxian model. In general, this means that a hybrid system - now understood in the general, that is, logical sense - sets the criteria of rule variation itself, in fact depending on its actually reached states. So there are no explicit meta-meta-rules, but only general, i.e., non-specific schemes which determine *that* it may come to a variation of the rules, that this can occur in different states, and with a different degree of radicalness. However, the question of *how* those schemes are put in concrete form remains an empirical one.

I would like to illustrate this with a little example: the "mechanistic" theories of pauperization oriented to Marx postulated that a revolution would occur whenever the degree of material pauperization exceeded a certain limit - of course they never said exactly how this limit could be determined decisively. By contrast, history seems to teach that there are no such absolute thresholds; on the contrary, people seem to tend to revolutionary changes just at times when the social circumstances have already been changed, that is, when the very idea of the possibility of change has become concrete because of the social development itself, as it were. Both the French and the Russian Revolution only took place when the old regime had already begun to carry out half-hearted reforms. In general, we can learn from the observation of various societies that the subjective readiness for social change is developed the furthest in places where the social rules have already become variable; in this respect, a

society like that of California, where Feyerabend's principle of "anything goes" seems to be a social constant, is surely an extreme in comparison to traditional agrarian societies.[67]

If these observations are correct, we can obviously speak of a positive feedback between the variability of social rules and circumstances and the subjective readiness of taking an active part in the enforcement of changes. However, this is only possible to a certain degree: the more variable the social rules are, and the more complex the own dynamics of the social system (the base system) become as a result, the more the system drifts off towards chaotic behavior. In turn, this causes a growing need for order, which fosters a social trend towards conservative behavior. The system decreases the variability of its rules or annuls them, which leads to a return to "smoothed out" trajectories and perhaps attractor states in the long term. Therefore, in order to start a development that is positive with respect to particular goals, the system has to increase its Wolfram class by a variation of its rules, as it were, at best up to class IV; in case it gets too close to the edge of chaos it withdraws into lower classes again. Of course, this does not necessarily have to happen fast enough, so the system can end up in chaos at least for a short time.

So in this respect the system produces both the readiness for changes as a criterion of the variation of rules, and conservative tendencies, including the withdrawal of rule variations, or possibly of rules themselves. Now we can understand the political institution of elections that offer the choice between conservative and progressive parties as a built-in mechanism that non-violently carries out the transformation of current tendencies of the mind into rule variation and rule stabilization respectively. Of course, this requires the existence of alternatives.

In section 2.4. we have already mentioned a formally quite similar problem, namely the development of limits of discrimination between different problem classes for learning and remembering systems. Following Holland (Holland et al., loc. cit.) I have put forward the plausible hypothesis that adaptive systems determine the corresponding discrimination parameters themselves in the course of their evolution, depending on their ability to comprehend their environment. It seems that this consideration is also true with regard to the "small enough d": it is a trivial fact that an adaptive system does not have an infinite layering of meta-levels to determine the size of d, but it keeps those parameters

---

[67] The peasants' revolts in the Middle Ages in Europe and other feudal advanced civilizations, or, e.g., the revolts of the weavers in Germany in the 19th century have of course demonstrated that people do strive for changes because of material sufferings. However, it is not possible to derive a general law from these examples. We will be dealing with this in more detail in chapter four.

variable, that is, dependent on its own evolution. Highly developed industrial societies have higher standards than purely agrarian societies. We shall see in section 3.3.3. that such an evolution of specific parameters - the so-called meta-parameters - seems to be a universal phenomenon-in connection with complex adaptive systems.

However, this model of rule variations does not release us completely from the problem of the external. We need to presuppose a certain "constitution" of psychic systems - no matter whether we regard them as subsystems among the social systems or as "structurally coupled" (Maturana) environments of the social systems. The assumption *that* psychic systems are in a relation of covariation and feed back with the changes of social states in the way that was outlined above can only be proven empirically and does not a priori follow from the properties of hybrid systems' rules. One might say that an anthropological condition has a hand in matters here, and conditions of that kind must always be presupposed. By determining specific criteria for SOZION and SOZGEN we have "just" shortened this condition for practical purposes.

It was mentioned above that SOZGEN already contains 3rd level rules. In order to come to a final clarification of our problem, we are going to take a closer look at the 3rd level rule - a meta-meta-rule - that controls the variation of the mutation rate:

In contrast to the switching on or off of the GA, the variation of the mutation rate does not take place on the basis of a fixed criterion. Rather, SOZGEN "decides for itself" on whether the mutation rate is increased, decreased, or kept constant. Technically, this is realized in the following way: the GA vector contains a component that is ternary coded, as in the case of the variation of the probability parameters; it varies the mutation rate or keeps it constant depending on the actual value. It is important to keep in mind that this component, which does not refer to the rules of the CA but to the GA, is subject to crossover and mutation just like the other components. However, the GA has an effect on itself. Thus, SOZGEN not only decides on specific rules of the CA, but also on a specific mutation rate, and does so, in fact, *without the use of external criteria.* So at first sight it looks as though SOZGEN is really able to set the criteria for the application of particular meta-rules itself - the alternative between infinite regress and externally preset criteria is resolved in favor of a complete self-control of the system which takes the criteria for its decisions exclusively from the concrete proceeding of its own dynamics.

It is obvious that this only transposes our problem of external criteria. SOZGEN needs at least one external predetermination, namely that it has to optimize at all, and in particular that it has to choose the best of its alternative possibilities. This is not necessarily a matter of course in the

case of real systems, which might as well prefer solutions that may not be the best, but demand the least possible change (see below, section 4.1.). Such a preference could explain, e.g., why changes that actually have been considered necessary by the system just do not happen; in particular, it would be an explanation for the continuous non-appearance of revolutions, which has always irritated Marxian theorists. As a result, it is indeed possible, like in the case of SOZGEN, to install the decisions on when, how radical, and how frequent meta-rules are brought into action mostly inside the system and thereby evade the inelegant external predetermination of absolute criteria; however, *something* has to be predetermined; neither real nor formal systems can fully control themselves *with regard to their variability* by their own efforts alone.

So maybe theoretical sociology, like theoretical physics, cannot do without constants, particularly when formal models like the ones described here are used. It is characteristic that Marx himself had to introduce a constant for his solution too, namely the "anthropological constant" of work (Habermas 1968), which keeps the development of the forces of production going and thus actually initiates the dynamics discussed above. This might be deplorable from the point of view of a "Kantian stance" (Kauffman), but one just can't have everything. In a similar context, Hofstadter (1985) has also pointed to the fact that any postulated and modeled variability requires a founding level which in itself must be kept constant *with regard to that variability*. In this case it is the anthropological level, and I suppose we have to be content with that.[68]

The above mentioned model of a universal grammar according to Chomsky is a comparable attempt to solve the problem. During its linguistic development the child needs criteria on whether and when it should change from too simple, and possibly incorrect, rule systems into extended and corrected ones, that is, criteria on when it should use the meta-rules for the variation of the already "implemented" linguistic base rules. Those criteria can only be biologically founded, which means that in addition to the usual concept, a universal grammar must also contain the deciding factors for the use of the meta-rules. After all, it seems to apply to all mental abilities that both their acquisition and their *variable*

---

[68] One might of course make an excursion into science fiction here: Maybe the progress of social technology in the ongoing project of changing the human species by means of genetic engineering through the species itself can cause a new variability on the anthropological level. Then the only founding level left would be that of the laws of nature, especially those of biochemistry. It is no accident, by the way, that theoretical physicists keep struggling with this problem too: How can the constants of nature be derived from the laws of nature? As far as I know, physicists did not succeed in finding a logically more satisfying solution yet either (see, e.g., Barrow 1994).

use point to the fact that they have to be modeled as hybrid systems whose "final" criteria are biologically founded.

3.3.3. The Parameterization of Rule Spaces: Variability Degrees and Meta-Parameters (together with Jörn Schmidt and Ralph Kier)

In sections 3.1.1. and 3.1.2. we have explicitly pointed to the importance of "control parameters" for dynamic systems, and also to the fact that they are all connected in a way that has not yet been clarified in detail. They can make actual system dynamics, in fact those of whole classes of systems, very transparent; in particular, it turns out that by classifying the parameters, the sets of rules of the various systems can be directly assigned to the classifications of the system behavior: specific sets of values of the rule systems' λ, P, Z, CF, or v parameters correspond to specific classes of complexity or Wolfram classes of state changes; as we could see from the analysis of the FP parameter, this also goes for stochastic systems. Hence, the degree of ordered behavior is directly determined by the control parameters, and can be significantly changed by varying them: Hofstadter (1986) once described this quite vividly in connection with another λ parameter as turning the λ–knob.

GAs have rarely been analyzed from this aspect, because usually it was only their "efficiency" that was of interest, that is, the quickness and precision with which they could reach a particular goal value. Incidentally, the same is true for hybrid systems. However, we could already see from the little experiments with the prisoner's dilemma that were described in section 3.2. that the genetic operators of crossover and mutation actually function as control parameters: a relatively high mutation rate - even a value of 0.1 was enough there - "destabilized" the system in that it caused repeated upward and downward fluctuations in the respective graphs of proceeding. Crossover played a part in this as well; if comparatively large segments are taken at once from one vector and put into the others, then this is not necessarily more efficient in comparison with smaller segments, but the system tends to become more unstable, that is, the trajectories rarely find a simple attractor.

If we try to analyze hybrid systems with regard to possible control parameters, accordingly, the main issue is obviously what influence the *variations of the rules of the base system have on its dynamics*. We have seen that only the system's states determine whether it has reached satisfactory system values; the example of SOZION showed that in itself the GA vectors and genetic operators were meaningless. If, as in this example, the base system is a CA or an equivalent system, the quest for possible control parameters can be understood as the quest for the

necessary conditions of the *hybrid* set of rules to cause the base system to show a particular adaptive behavior. To put it more precisely: control parameters of hybrid systems - if they do exist - can give information about whether and to what extent the adaptive behavior of the hybrid systems depends on the values of the "hybrid control parameter".

It is evident that such parameters belong to a different logical level than the well-known parameters of CA and BN. In the case of hybrid systems such as SOZION or SOZGEN, one cannot speak of "the" Wolfram class to which the systems belong, because of their sets of rules or the control parameters that characterize them. But there are external goal values - the environmental requirements - to which the system development orients, which can finally lead to a variation of the λ or Z parameter if the meta-system - in our case the GA - is frequently applied. A more radical rule variation than in the case of SOZGEN or SOZION can in practice cause the developing of any rule system and thereby - at least for particular time intervals - make, e.g., the λ value fluctuate between the classes I and IV. There will be an example of this later.

Some minor - unintentional - experiences with SOZION can show this more clearly. When we constructed the set of rules we did not explicitly consider the dynamics of the CA, because we intended a hybrid system from the start. In later experiments without the GA, we noticed that the original rules of the CA unambiguously placed it into class II. After about 15 runs, the automaton often reached an attractor state of the period 0, that is, there usually was a homogeneous society that exclusively consisted of members of the highest class; other simple attractors are also possible, though rather exceptional. The high social differentiation and the high complexity of the whole system were only achieved with the GA, which made the CA behave as if it belonged to class IV or III. When the GA was switched off, the CA immediately fell back into a behavior of class I or II. By the way, the CA of SOZGEN is in a higher class of complexity because the rules of population cannot be switched off there.

As a consequence, we have to search for a "hybrid" control parameter on the level of the meta-system. Since the most interesting aspect of the behavior of hybrid systems is their ability to orient as quickly and successfully as possible to specific environmental requirements, it should be the function of a meta-parameter to serve as a measure for just this ability. Thus, we should determine how variable the system is *on the rule level of the base system*, and then correlate this variability with its effectiveness of goal achievement. Once again we would like to emphasize that this results from the fact that on the one hand a hybrid system's base system is usually not in one specific class of complexity, but instead it may "jump" from class to class; and on the other hand, the variability of the base rules only make sense to the system as a condition

of adaptation because it has to achieve specific goals. Variability for its own sake is of course pure nonsense. As a result, hybrid control parameters or *meta-parameters*, as we shall call them in the following, measure the effectiveness of adaptation, which, after all, is also the criterion for the characterization of different types of GAs: it can be the amount of time that an adaptive system needs to realize specific system values to come up to specific environmental requirements; it can also be the realizable system value itself - the higher, the more effective; finally, it can be a combination of both. Usually, what matters to an adaptive system is not the attainment of especially good system values; in most cases, it is more important to reach fairly acceptable values in the available period of time. That is why we used the time that was needed for particular optimization processes as the criterion in our experiments; however, in several cases we used in addition the level of the generally reached system value.

On the basis of our previous considerations and examples it should have become intuitively clear already that there are many different ways of changing the rules of a base system by meta-rules. Our present experiments with meta-parameters, whose most important results will now be described, show that there are also different meta-parameters, which have accordingly different effects on the behavior of the base system. On a purely qualitative-phenomenological level we presently distinguish between meta-parameters that (a) change the rule system in a specific *quantitative* way; and (b) do that in a specific *qualitative* way. This will be explained in the following:

(a): if we look at the total number of rules of a base system, we can determine the following quantitative degrees regarding their change: the first one concerns the question of how *frequently* rule changes have been carried out within a particular period. In the case of our formal systems, this degree is made up of the proportion of the total number of runs of the base system and the number of rule changes. Let us call this degree h (for *Häufigkeit* = frequency), let the number of runs be n and the number of rule changes m. Then it is evident that

$$h = m/n. \qquad (3.1.)$$

Second, we have to investigate *how many rules* are kept variable by the system, i.e., how many rules are kept constant by the system and, accordingly, how many are allowed to be varied. We have called this meta-parameter "the principle of the sacred cows", and it is easy to see why: every system, and in particular every social system has certain rules that are not disposable, no matter how high the pressure to change is for the system. Among them are especially the rules that concern the

system's identity, e.g., in social systems rules that characterize specific traditions, religious and/or cultural norms, and general values from the system's point of view. In Western democracies this is the case, e.g., for the principles of Human Rights.[69] Let us call the degree of rules that are allowed to be varied r (for "radicalness" of rule changes), let the total number of base rules be k and the number of variable rules j. Then

$$r = j/k. \tag{3.2.}$$

Third, we have to ask *how many rules at a time can be changed by an operation of the meta-rules*. This is not the same degree as r, which can be illustrated with the crossover operator of the GA: in SOZION all probability parameters were allowed to be varied in principle; if we only consider these as rules for the sake of simplicity, which is of course actually incorrect, then SOZION has the value r = 1 (by the way, h = 0.5 in SOZION, which means that there is a rule variation after every second step of the CA). However, only about 20% of the mobility parameters are actually varied at one single rule change because of the segment's length at the crossover plus the mutation rate. Let us call this degree g (for *Größe* = size of the varied rule segment), let the total number of base rules again be k and the number of rules that are varied at once i, then

$$g = i/k. \tag{3.3.}$$

Thus, in SOZION g = 0.2.

Finally, we have to analyze *how subtle* the rules are changed. What this means is, of course, that the rules can be changed more or less slightly. That sounds rather like a qualitative degree; for rule variations that are carried out in CAs like SOZION and also in neural nets, the degree can easily be quantified, though: in both cases the important variations take place as changes of numerical values in a matrix. It is obvious that this can happen with varied subtlety. SOZION, e.g., always changes its matrix values - the probability parameters - by means of increasing or decreasing them by a maximum of 0.1. In comparison, the weight matrices of hybrid interactive neural nets are changed much more radically by the crossover; for example, there can be value changes from 0.1 to 0.9 at one variation step. So in this case s is much lower or, in other words, the "robustness" of the changes is much greater. Thus, s can be described as the maximum degree of increase or decrease of such matrix values. If we reduce the scope of the generally possible variation

---

[69] In the long run, however, all rules are disposable or at least they can be slightly modified if need be, which can be seen, for instance, from the changes in connection with the right of political asylum that took place in almost every Western European society. The above determination always has to be seen in connection with specific observation periods.

of these rule values to an interval between 0 and 1 for the sake of technical simplicity, and if we call the maximum value change that can be realized by a single application of the meta-rules t, then

$$s = 1\text{-}t. \qquad (3.4.)$$

Thus, in SOZION s = 0.9; in the above mentioned hybrid interactive networks we have s = 0, since the values can change from 0 to 1 and vice versa.

We could easily think of further quantitative degrees as meta-parameters, since Kauffman's above quoted remark that there is more to be discovered can surely be applied here too. However, we think the meta-parameters mentioned above are the important ones whose properties should be analyzed.

(b) The above distinction between quantitative and qualitative rule changes is doubtless somewhat artificial, because, e.g., s does indeed measure qualities of rule variations as well, and the meta-parameter μ, which will be our next subject, is naturally also defined in a quantitative way - otherwise it would not be a parameter in the strict sense. The difference between the meta-parameters (a) and (b) can be marked a bit more precisely: the meta-parameters of type (a) have been defined as specific quantitative relations with regard to the totality of base rules and/or the observation period or the number of runs. We did not take into consideration yet that the meta-rules can have different effects on the dynamics of the whole system, though the meta-parameters of type (a) have the same value. As a consequence, a "qualitative" meta-parameter has to be explicitly determined as a variation of the dynamics, that is, as a variation of the base system's class of complexity.

Before we realize this formally, we would like to describe briefly the problem with the example of SOZION: we mentioned before that SOZION is a hybrid system whose CA, without the GA, belongs to Wolfram class II, but merely reaches a point attractor in most cases. Even a variation of the probability parameters and/or the switching off of single rules does not change this in principle, which was already analyzed with the example of the FP parameter in section 3.1.; the complexity of the total dynamics of the *hybrid system* SOZION is in fact only due to the GA. So, in spite of high values of r and s, the dynamics of SOZION's CA are usually not significantly changed.

Such a change can easily be achieved by introducing additional rule changes. The following is an example of a standard rule that is usually not changed by the GA: if the (arithmetic) mean of the states of a cell's neighboring cells has a value higher than or equal to the state of the cell itself, then the cell "moves up" one class (according to a particular

probability); if the mean value is lower, it moves down one class. If these rules are changed so that just the opposite is true in the case of higher classes, the CA immediately turns into class II and almost always reaches attractors of the period 1 or higher; thus, this rule variation realizes a change regarding the class of complexity.

By generalizing this example we come to the following assumption: it seems to be possible to classify rule changes according to their capability to cause a significant change to the base system's dynamics. As we have seen before, these dynamics can be characterized by the values of our control parameters, and thus a corresponding meta-parameter could be defined by its capability of giving a particular degree for the variation of one or more control parameters.

There are several control parameters and until now we could only describe their connection by the merely qualitative considerations of section 3.1.2. Strictly speaking, one meta-parameter would be needed for each control parameter until more exact connections are found. On the other hand, the single control parameters all are different measures of the particular dimensions of inequality the rules of the system will produce. That is why we decided to experiment with only one control parameter first, namely with P and its logical equivalent λ, respectively. Since the accompanying meta-parameter operates *on* P and λ respectively, we have called it μ. Hence, μ generally measures the changes of P (and λ respectively) of a base rule system, which is described by the difference between the values of P before and after the rule variation.

The problem that arises from such meta-parameters, however, is that they cannot be determined irrespective of the number of operations of the meta-rules. This can be illustrated with a little example: like all control parameters, P varies within a fixed interval; in the case of P, this interval is between 0.5 and 1. Let us assume that we have a rule system of a CA with $P = 0.8$ at the beginning, and that the meta-rules work in the way of the different cell states becoming more and more uniform. As a consequence, there will be an almost equal distribution sooner or later, so that the P value of our CA will tend towards 0.5 and remain there - one could also say that the CA's rule system de facto reaches an attractor as well through the iterated application of the meta-rules, a meta-attractor, so to speak (see below). Of course, this continually decreases the difference between the various P values as well, and with that μ varies in dependence on the number of iterations (of the meta-rules): the more often they take place, that is, the higher our system's h value is, and the smaller is the corresponding value of μ.

Therefore, μ can only be defined as a mean value for a particular observation period or a particular number of runs - the same goes for h. Thus, if m is again the number of rule variations, and if

$$\mu_i = |(P_i - P_{i+1})|,$$

that is, the change of the P value as a result of applying the meta-rules i times, then

$$\mu = \Sigma\mu_i/m, \tag{3.5.}$$

where i runs from 1 to m.

The simple variations of the probability parameters in SOZION obviously bring about $\mu = 0$ in this program; if the previously described additional rule variation is iterated so that the conditions are inverted each time, we get μ = approx. 0.2 for this hybrid system.

So far, our considerations have always referred to hybrid systems which consist of two logical levels: base system and meta-system. Of course, this is not a necessary restriction; the formal system SOZGEN already has three logical levels: the CA as the base system; the GA as the meta-system with the genetic operators and the already mentioned meta-rules; and the additional rules that control the switching on or off of the GA and the variation of the mutation rate. The latter are obviously 3rd level rules, because they refer to the meta-system. Therefore, hybrid systems consist of *at least* two logical levels. From these assumptions we can draw the following formal characterization of the variability of hybrid - meta-dynamic - systems:

If we combine the different meta-parameters of type (a) in a "total parameter" κ for the sake of clarity, and if we describe the number of logical levels which the system consists of as ν, then a meta-dynamic system's variability V can be characterized by

$$V = (\mu,\kappa,\nu) \tag{3.6.}$$

At present, the question of whether and how the single components of V are connected mathematically is completely open; this will be a subject of further research. (The connection between stochastic parameters like FP and the mentioned meta-parameters is equally unsolved.) What is important here is that V does not have to be a constant, since the variability of a hybrid system can of course change as well. We can see a similarity to our analysis of the characteristics of complexity and emergence here: a quantitative determination of V can only be valid for a specific observation period.

Before we present some preliminary results regarding the meta-parameters in question, we would like to emphasize another aspect of adaptive systems, by which the difference between systems that are "merely" self-organizing and adaptive systems can be elucidated once more. It concerns the fact that there seems to be yet another difference

between the various meta-parameters apart from the ones we already mentioned:

We know that in the case of CAs and BNs the specific values of the control parameters cause the corresponding system dynamics. Thus, the question of research in connection with the analysis of these parameters is always this: which fixed parameter values result in which unequivocal system dynamics? The corresponding question with regard to the meta-parameters would be: which values of the meta-parameters produce which optimization processes and, in particular, which values of the meta-parameters produce the best optimization behavior - the "best adaptation"? Of course, these questions are important to every system.

Interestingly enough, in contrast to purely self-organizing systems, we have not only one, but actually two questions here: we can indeed analyze which optimization processes arise from particular fixed values of the meta-parameters; however, for some meta-parameters, and perhaps even most of them, this does not provide an answer regarding the best optimization processes. It is much more likely that a system can cope best with its environmental requirements if it *keeps some meta-parameters variable* by differentiating between rough and fine adjustments, as it were. What this means can be explained with the example of the delta rule for artificial neural nets that was introduced in 2.3.; in a somewhat simplified form the rule looks like this:

$$w_{ij}(t+1) = w_{ij}(t) + \delta,$$

where δ is the "error" at the time t, that is, the difference between the actually realized state and the goal state. Since the realized states constantly change, namely towards the goal state, δ changes too: at the beginning of the learning process δ is relatively great; it continually decreases with the improvement on learning. Thus, the system realizes its learning processes or optimization processes by keeping a particular parameter variable: the closer the system approaches the goal values, the finer is the adjustment between actual value and goal value, because the necessary corrections can (and must) become smaller and smaller.

It is evident that δ is a meta-parameter too, which determines the weights' range of variation. If we generalize these results, which we will encounter again in section 4.3. in a different form, we can explain them - at least hypothetically - in the following way: the most favorable behavior for a system in an optimization process that can be characterized by the iteration of specific rule and parameter variations is to readjust constantly particular meta-parameters, in fact by performing great changes at the beginning of the process and then constantly carrying out smaller variations while the approximation to the given goal states gets more and more refined. There certainly are meta-parameters to

which this does not apply, i.e., where there are particular fixed values that make an optimal adaptation process; we shall present an example of that as well. However, many meta-parameters, and probably most of them, can be characterized by the above assumption.

Now, from the point of view of our hypothesis, let us take a look at the above meta-parameters, which we have studied with different hybrid CAs (with GAs) and, together with Christina Stoica, with hybrid interactive networks. In the case of h, the results are unambiguous, in fact in the way we expected them to be: the optimization success of the adaptive systems that we analyzed is greater the greater h is, in other words: the more often a system carries out the respective variations, the better is its optimization. This result is not surprising per se, since it is immediately plausible: life punishes those who do not carry out reforms often enough.

These results alone do not say very much, since the success of changes fundamentally depends on how thoroughly and how extensively they are carried out. We shall see indeed that frequent variations alone are often insufficient to attain favorable values in acceptable time. If the changes do not take place "in the right way", the optimization processes can get stuck, as it were, that is, they do not get beyond a particular actual value which need not be sufficient in itself. We already pointed to that in connection with our general reflections on the GA. We notice, however, especially when we look at it from the above mentioned aspect of fine and rough adjustment, that without a minimum size of h - depending on the difficulty of the optimization problem and the size of the system itself of course - there can be no acceptable optimization at all.

Our present results indicate that g really is a parameter that can be varied best by the system itself. In the case of SOZION it turned out that the system generally did not get beyond a distance of d = 0.08 from given environment vectors for g = 0.2. However, this distance was attained rather quickly. When g was continually reduced, it turned out that on the one hand the attained results were getting better and better, but on the other hand SOZION needed more and more time (measured in the number of runs) just to attain the value d = 0.08. In the end, SOZION often attained distance values of d = 0.000, i.e., the absolute optimum, with g = approx. 0.012; as a consequence, the corresponding optimization processes took a longer time of course. Thus, the price of favorable results is that the system has to expend more time, and consequently has to carry out changes much more often. If h is small, a small g value would consequently bring only suboptimal results.

. Obviously, the system would be best off if g could be kept variable: the system should start with a great value of g to minimize the necessary time, and lower g when there is a notable approximation to the goal values. Of course, these are only qualitative results, since it would require a long-term series of experiments to find out the exact values for the variation of g. At any rate, the present results clearly suggest that g - like δ - is a meta-parameter which should be varied by the system itself. This is mainly because real systems can never spend unlimited time and energy on reforms just to attain a good result in the long run, since, as Keynes once put it in a similar context: "In the long run we are all dead." On the other hand, "good is just not good enough" sometimes, to use an advertising cliché. A variability of g can solve this problem.

In comparison, the results regarding s are not so unambiguous. We pointed out that SOZION has a relatively high s value which enables it to attain regularly favorable actual values - in dependence on g. However, we could see that with high values of g the problem arises of the system getting stuck in suboptimal solutions. In this case, not only a reduction of g but also a reduction of s can help, which will be shown in detail in section 4.2. We must consider, though, that low values of s have the unpleasant side effect that the system might as well leave favorable values already reached, as can be seen from hybrid interactive networks. Thus, a forcible change of the rules might get the system out of blind alleys, which often really works by means of mutation in a GA; but it might just as well have a counterproductive effect, because after the change the system can easily come to even worse solutions. Again, a variability of the meta-parameter seems to suit the system best, though it is not as simple as in the case of g and δ: in principle, a high value of s is best for the system; but if the system does not get any further with these values there needs to be a reduction of s, which has to be withdrawn after the changes have taken place. Hence, the variation of s would be an emergency measure that the system reserves for evolutionary blind alleys. In 4.2., we are going to describe *how* social systems and those comparable with them carry out such variations of s, namely by means of self-modeling and the mind.

In the case of μ things are quite unambiguous again. Here, unlike in the above described cases, we laid out our experiments in such a way that the GA had to determine the most favorable values of μ itself. It turned out again and again that in the case of random generations of CA rules and their variations, the closer the system got to the given goal values, the more μ was reduced. If the system happened to reach its goal values, the GA set μ to $\mu = 0$ and so the CA rules did not change anymore with regard to their P value. As for the CA, the given goal values' meaning is of course that it has to attain some distribution of the different cell states

over the entire grid; therefore, we can also understand optimization in terms of the GA generating a CA rule system with a P or λ value that matches the required distribution in the grid. This is exactly what happens: the GA leads the CA rule system closer and closer to the required P values by constantly reducing μ.

This outcome is remarkable in many respects: first, it turns out that μ clearly belongs to the meta-parameters that need permanent rough and fine readjustment in dependence on their distance to the goal value - a point that was not necessarily foreseeable by intuition. Hence, one cannot speak of "the" most favorable values of μ. Instead, an adaptive system is best off if it changes its actual complexity class as drastically as possible, if necessary, at the beginning of its optimization process, and subsequently makes continually slighter changes of its own complexity. Second, the solution that the hybrid system carries out, namely to generate rule systems for itself that have a P value that guarantees a solution of the optimization problem, may be extremely plausible; however, this does not necessarily have to be the case, since other solutions to this problems are possible as well. The fact that our hybrid systems always offer this type of solution demonstrates the relevance of the control parameters for self-organizing dynamics from a completely different point of view, namely that of hybrid adaptive systems. Metaphorically speaking, our hybrid systems have rediscovered the P parameter for themselves, as it were.

Finally, these results show a strong point of the genetic algorithm, which we noticed in other contexts as well: the behavior and results that hybrid systems show are often well known from real life. It seems that these formal systems are not at all constructed at random, which is how they appear at first sight. They rather correspond to real processes in an amazing way sometimes. After all, the error δ has been entered by human constructors into the delta rule on the basis of Hebb's model; in our case, the formal system has discovered the necessity to introduce an appropriately variable meta-parameter and the relevance of the control parameters of its own accord. In section 4.3., we are going to come back to that again.

By the way, the determination and variation, respectively, of μ by the GA constitutes in itself the introduction of a third logical level into our hybrid systems, because the system itself varies the variations of its base rules, that is, it operates upon the meta-rules. Since these "meta-variations" obviously depend on the states of the base system, we can derive an example for an already mentioned case here: the logical layering of different rule levels must not be understood as a one-sided hierarchical relation - unlike the layering of levels in logical semantics

(see above, section 2.5). Not only is there a variation of the base rules "from above", but the results of these variations on the level of the base systems also determine the meta-variations on levels 2 and 3 again. There is nothing paradoxical about that, it rather draws our attention to the fact that not only social and psychic, but also formal systems can consistently handle a continual interdependence of different logical levels (see below). When Hofstadter and Dennett say in our opening quotation that "a clear understanding of that interdependence is probably a good ways off yet", then the systematic construction and analysis of hybrid systems can obviously help us to make considerable progress on this way towards this sort of understanding.

So far, we have not said anything about r, although we can see important characteristics of the variability of adaptive systems here as well. In section 2.4. we already mentioned that from a systems theoretical point of view memory is just the determination of specific rules, and thus a reduction of r. The reason why we did not deal with r before is that our present results regarding r are not unambiguous. Christina Stoica has submitted results (Stoica 1997) that show that a reduction of r means an improvement of the optimization behavior in hybrid interactive networks; to put this point more precisely, in connection with the optimization of special hybrid interactive networks it turned out that the best optimization performances could be achieved with values of r between 0.5 and 0.6. It seems that these systems follow the maxim: "Change as much as necessary and as little as possible." We obtained similar results with our hybrid CAs, which could not, however, always be confirmed.[70]

There seem to be two reasons for this remaining uncertainty: on the one hand, r has probably been defined too roughly, since it is an implicit part of the definition that all rules are equal, that is, they have the same effect on the system. But if we think about it we can easily see that this is usually not the case, either in formal or in social systems. Thus, a precise definition of r should consider the respective relevance of each rule as well. This can in fact be realized quite easily for CAs and similar systems: when we introduced the λ and P parameters in section 3.1., it became clear that the dynamics of a system fundamentally depends on how many transitions of single cell states into others are realized by the rules (in stochastic automata also: the probability-frequency with which the transitions into the different cell states are generated by the rules). On the basis of these considerations it suggests itself to classify the

[70] The formally comparable studies by Carley (1997) regarding the adaptivity of simulated companies that operate on the optimization principle of simulated annealing produced similar results. Carley does not analyze these questions systematically though, and besides we have also obtained different results. The unambiguity of Carley's results might be due to the chosen method of simulated annealing.

importance of single rules for a system according to the number of transitions that are realized by the different rules and their probability, respectively. Thus, if the number of transitions that can be realized by the rule R is n, and the number of transitions that can be generated by the rule R´ is m, then R is more important to the system than R´ if and only if $n > m$.[71] Correspondingly, we experimented with the determination of rules that were classified according to their importance, that is, rules were divided into "relevance classes", and the variation of r was done in such a way that the same number of rules from each relevance class was determined or allowed to be varied. The results regarding the hybrid interactive networks could be confirmed on special conditions, but, again, not in every case. It seems that a consideration of the rule relevance is not quite enough yet.

On the other hand, our reflections on the memory of systems showed that the importance of rule determination, that is, the reduction of r, often lies in the fact that the system just needs to remember instead of having to carry out new optimization processes. In our present experiments this has not been taken into account yet, which makes it another desideratum.

However, the assumption suggests itself that r is also a variable parameter which is adjusted by the system as required. The considerations on the subject of learning in chapter 2 showed that the newer and greater a problem is for the system, the more rules have to be changed. In particular, memory, i.e.; rule constancy, becomes counterproductive when the system is prevented from grasping radically new situations. Once again we can point to Bateson's learning III here. In this case, radical increases of r are obviously reasonable and necessary. If, on the other hand, the problems are not particularly unusual and the system has already come quite close to the goal values, then low values of r are most favorable so that the system does not have to carry out too radical changes.

In addition, it can be assumed that specific values of r also depend on the total size of the system, that is, its number of rules: if a system has many rules in total, then there is still a great number of rules available for variation - considered in the absolute - which is sufficient to achieve favorable solutions even if the r values are at about 50% (since r only measures proportional values). In the case of relatively small systems the absolute number of variable rules might not be enough when there are low values of r. In fact, this has already been verified for small rule

---

[71] In logically one-dimensional systems with $v = 1$ it should be taken into consideration that at particular initial states not all states can be generated by one rule or one rule system in general. Therefore, this definition of rule relevance is always about the possibility of the generation of states by a rule *in principle*.

systems: in a simplified version of SOZION where a maximum of 10 rule components could be changed it turned out quite unambiguously that the optimization success directly depends on how many rules are kept variable: the greater r is, the greater is the optimization success.[72] Again, there is still a lot of research to be done here.

It seems to be possible in principle to characterize hybrid systems formally in the same way as is possible with self-organizing systems. Of course, things are more complicated here, as was to be expected. Adaptive, i.e., metadynamic systems simply have at least one more logical rule level than purely self-organizing systems. Therefore, it makes sense to make some conclusive systematic reflections on the whole matter.

(a) The classification of a dynamic base system B into a specific class of complexity means that its dynamics is characterized by one or more simple attractors of a finite period, or by strange attractors. However, this says something about B's concrete trajectory only if we also know the state of B from which the trajectory starts out at the time of our observation. In the case of systems of class IV it can easily occur that their trajectories end in an attractor of the period 0 after a short time, so that they seem to belong to class I at first. Everyone that has experimented with the "Game of Life" (class IV) has probably experienced this. Wuensche and Lesser's "atlas" (loc. cit.) of the basins of attraction of different CAs demonstrates quite clearly that the single trajectories depend on the rules *and* the initial states. Hence, the fact that a base system belongs, e.g., to a high Wolfram class only indicates that the system may possibly achieve long periods, which is impossible in a lower class. On the other hand, the knowledge of the future "history" of the system always requires both types of information.

This is different in the case of hybrid systems. A system like SOZION can make itself independent of practically any initial state of B by the fact that the meta-parameters that are combined in κ have values that entail a high variability. As we said before, without rule variation the homogeneous initial state of SOZION always leads the CA into a final state that is a point attractor, and often consists of members of the highest class only. On the other hand, practically any state that is actually possible for SOZION can be realized by suitable choice or variation of the environment vector, because of its relatively high variability (the number of possible states is 10^{4000}). In other words: the higher the variability is of a hybrid system H, the more independent is the base system B of any single states with respect to its future trajectory (and, of course, the more difficult is the anticipation of that trajectory).

---

[72] These experiments were carried out by Tanja Böhnke.

If we interpret rule variations with regard to their "content", namely as a system's attempts for reforms, we know that in the case of social and other hybrid systems like, e.g., the mind, reforms are usually arduous and are only carried out if absolutely necessary. As a consequence, a hybrid system obviously has a not insignificant choice: if it wants to carry out as few reforms as possible in order to carry on with a constant rule system of B as long as possible, and if it is in a low class of complexity, then a high value of μ is required - the reforms have to be carried out quickly but radically to give B the necessary dynamic complexity. As we know, in the case of rule constancy B can reach many different states only if it reaches a high class of complexity. If a system shies away from radical changes of its dynamics, then reforms have to be carried out continually with a high value of h; this might be the standard method in democracies, since radical changes are only rarely asked of the electors - justified or not.

(b) We also have to clarify the meaning of the attractor concept for variable, and thus hybrid, systems. In section 3.1. the attractor concept was exclusively used for self-organizing systems, which is where its meaning can be directly seen: a rule system f, that is, a system function, generates a trajectory that either realizes specific attractor states, or, if it is the trajectory of a chaotic system, can at best be characterized by strange attractors, depending on the system's initial state. This results in the sufficiently analyzed possibility to describe the system dynamics with different classes of complexity, and so forth.

This is not so simple in the case of variable systems. The rule variations - especially if there are high values of κ and μ - can constantly force new trajectories and change the base system's class of complexity with the same constancy, as it were. In particular, we already pointed to the fact that variable systems can make themselves independent of practically any *specific* initial state; this naturally goes for any attractor state too: if a base system has once reached an attractor, a slight rule variation is enough sometimes to take the system away from the attractor again, which is in fact often the object of the variation. Thus, variable systems have no attractor at all in principle, as long as we understand attractors according to the definition in chapter 2 as states which a system either does not leave anymore - period zero - or to which it continually returns after a particular number of states. It seems that the concept of the attractor is unimportant as regards hybrid systems, however fruitful it has proven to be in other cases.

In fact, things are not that simple of course. As we can see from the example of a hybrid system such as SOZGEN, especially adaptive systems do not permanently carry out rule variations either. On the

contrary, most systems tend to preserve the new circumstances - and rules - after they have made a change. We can certainly say that adaptive systems are generally conservative: they only change if they must, and afterwards they like to keep things the way they are. Therefore, it makes sense to analyze any hybrid system with regard to what kind of dynamics and which attractors the base system has after the process of rule variation and subsequent rule constancy - that is, in which class of complexity it is. Undoubtedly, any system is interested in quickly reaching an attractor of the period 0, which is not exactly easy, however: if the hybrid system adjusts its μ-value itself, the base system will be exactly in the class of complexity that it needed for realizing *one* favorable state after it has finished its rule changes. However, the trajectory of a high class of complexity usually leads away from that state quite quickly, if it should not happen to be a point attractor. We shall see in section 4.1. how hybrid systems deal with this problem.

However, distant from its original meaning, the attractor concept can be understood in yet another way in the case of hybrid systems. If we see the rule system of a base system B as itself a dynamical system - a second order system, so to speak - with the meta-rules as its rules of interaction, then we can look for the dynamics of this system, too, on that second order, and especially for its attractors. It is plain to see that either the rules of B are permanently changed by the iterated application of meta-rules without the original rules of B ever coming into effect again, or the original rules of B are realized again after some time. In the latter case we would have attractors for the rule system of B; by the way, if the meta-rules are deterministic, they even have to have attractors, as we have seen, provided they do not continually add new rules (see below, section 4.5). In other words: all our considerations of section 3.1 are directly transferable to our rule system and its meta-rules; the meta-rules' attractors are then simply meta-attractors of the whole system H. Accordingly, we could classify the meta-rules according to their attractors. Thus, a hybrid system can have attractors of two levels at the same time in principle, namely if on the one hand the meta-rules have produced a point attractor on the rule level, and on the other hand the base rules that are being kept constant by that take the base system to an attractor. However, these cases are probably rather rare, though one often gets the impression from observing political reforms that they work exactly that way: reforms (= rule changes) are permanently carried out without really changing anything in the dynamics of the system. Rules and meta-rules have possibly each reached a point attractor in such cases.

Accordingly, the remarks on the adaptive dynamics of biological evolution in section 3.2. have to be stated more precisely now: if we interpret genomes as base rules that constitute the base system B - the

phenotype - and if we understand the biological mechanisms of mutation, selection, and crossover as meta-rules, then a stable species can as well be understood as a hybrid system in which both meta-and base attractors have been realized. Although both meta-rules - the rules of biological evolution - and base rules - the "epigenetic" rules of the phenotype's ontogenesis - are constantly iterated, the species does not change.

It is certainly an interesting question whether these meta-attractors have the function of meta-parameters themselves, that is, whether they influence the adaptive behavior of H themselves. Intuitively, there is a lot to be said for this assumption, since in the case of a relatively great h it makes a difference whether the meta-rules have a (meta) attractor or not, that is, whether the rule variations get stuck as it were - which frequently happens with GAs - or if they constantly lead to new rule systems. This question, like many others, has to be left here as a subject of further research.

Let us become a bit more sociological again at the end of these rather formal considerations. By introducing concepts like "meta-rules", "meta-parameters", and now even "meta-attractors" we could hopefully show that formal metadynamic systems can be characterized and analyzed mathematically-experimentally with them. Thus, those strange looking concepts make sense on the level of formal systems. However, this does not say whether social systems have such characteristics too. If that is not the case, there would be the extremely serious problem concerning the validity of our formal models and the results from them.

However, there can be no doubt that practically every social system has quite unambiguous and usually easily identifiable meta-rules - the rest follows from that. Any political system has a lot of procedural rules that say exactly how the rules in force - the base rules - are to be changed. In societies like the Western democracies this is primarily linked with the competencies and standing orders of parliaments and administrations. There are 3rd level rules as well, namely rules for the variation of meta-rules: for example, a parliament can change the constitution to the effect that its own meta-rules for the variation of rules of social interaction are changed - any new balancing of competencies between the Federal Government and the Federal States in the Federal Republic of Germany is an example for that. Luhmann's famous remark that in those societies legitimations are achieved by procedure illustrates this quite clearly: the constancy of social rules as well as their variation is legitimate (and legal) if and only if this is based on generally accepted meta-rules. Societies are systems of social rules and their resulting interactions. Furthermore, they are systems with relevant meta-rules and

partly even meta-meta-rules. These are in principle well-known facts, though not usually stated in this terminology.

Of course, this goes for any social area, which should be easily understandable. Even social systems that seem rather informal, like children's play groups, have both interaction rules and meta-rules: Schwartzman (1978), following Bateson, describes very vividly how children communicate in a highly regulated way on whether, and how, certain rules of the game are to be changed, which new role allocations are to be carried out, etc.[73] We could easily find more examples, and what they teach is that meta-rules are obviously no esoteric invention of theorists who are far removed from social realities, but an integral part of social systems which they and social actors always have at their disposal as a matter of course. Obviously, this happens within the system without paradoxes and through selfreferentiality, which adds some precision to this woolly concept. However, this leads us to another section.

3.3.4. The Mathematics of Selfreferentiality

As we could see from the example of the variability degree V, hybrid systems cannot be classified as easily as can pure, self-organizing systems. The reason for that is, of course, that in vertically coupled hybrid systems we always have various sets of rules plus an orientation to external criteria, each of which has to be taken into account both separately and with regard to their interdependence. Hence, it comes as no surprise that things are a bit more complicated than in the case of the rule systems of BNs and CAs. On the other hand, hybrid systems offer possibilities for the formal modeling of problems that were for a long time not thought to be formally ascertainable. I have already pointed to that in general in section 2.5; thus, the analysis of hybrid systems presented here, which is certainly not completed yet, can be understood as an attempt to put the general considerations of 2.5 in more precise and concrete form. If we apply our sketched algorithms and formalisms to "real" systems, we obviously have a modeling of what has been analyzed in sociological systems theory as well, at least since Luhmann, namely the problem of the selfreferentiality of complex systems. Let us look at that problem a bit more closely.

It is known since the antinomies in logic and set theory that selfreferentiality can easily lead into paradoxes, that is, into a situation in which a system oscillates between two equally possible and therefore equally impossible positions, just like Buridan's Ass. This is usually

---

[73] I owe this hint to Helga Andresen (J.K.).

illustrated with the semantic paradox "This sentence is false." [74] The system has to get rid of the paradox in such a situation ("*entparadoxisieren*", to use a handy term from Luhmann). Pure self-organization does not lead into paradoxes, since the system constantly follows the same rules; paradoxical situations can only arise if the system is, e.g., confronted with an alternative of two equally favorable or equally unfavorable possibilities; in particular, a situation might occur where the following of different action possibilities leads to a negation of the actual intention and/or a worsening of the original situation. What shall the system do now to avoid starving in front of completely similar bundles of hay like Buridan's Ass?

Our reflections on hybrid systems show that there are indeed possibilities for a system with sufficiently high values of κ or also μ to turn to: it can try to solve the paradox by varying its rules; it can even experiment with rule extensions; and above all it can try to think of hypothetical paths of development by self-modeling to make its developmental capacity more effective. There will be some examples given in the following chapter. However, we already see that, on the whole, systems with the extensiveness of vertically coupled hybrid systems do not only have the problem of getting into paradoxes of selfreferentiality, but also have many possibilities to free themselves of them in a non-coincidental way. In particular, they may be able to realize that the situation is not paradoxical at all - it just looks as if it were (cf. note 74). According to Luhmann, this is only possible for an observer of the systems, and this is obviously particularly realizable in the case of our hybrid systems: on a meta-level, the system can observe itself quite well, namely, e.g., as a coding of the base system rules. Of course it can usually only observe its base part, but this is actually the one that is in the (seeming or real) paradoxical situation, and not the whole system H; of course, nothing prevents our system from entering even a third level in

---

[74] It seems to be an ineradicable fact that this classical paradox is still described with the example of Epimenides the Cretan (e.g. Kneer/Nassehi loc. cit., 105). "Epimenides said that all Cretans are liars". The explanation goes like this: if Epimenides is telling the truth he is lying and his statement is false, therefore he is not telling the truth. If he is lying then all Cretans tell the truth so he is telling the truth. Unfortunately, one overlooks the fact that the only conclusion that can be drawn from Epimenides' lie is that *there are Cretans that are not lying.* We can not conclude that Epimenides himself is not lying, since the sentence "All Cretans are not liars" does*not* follow from his lie. Thus, this is not in fact a paradox at all, but a proposition from which we may conclude that Epimenides was lying - which is empirically quite probable by the way. Some knowledge of the predicate calculus would probably do even sociological systems theorists some good.

order to analyze whether the - real or seeming - paradox is caused by the meta-rules and meta-parameters.

As a result, the formalisms of hybrid systems represent something that is actually not possible according to the logical semantics' bans on selfreferentiality (Tarski 1956), namely a mathematics of selfreferentiality. This is possible because of the fact that was analyzed in section 2.5, namely that these systems operate in real physical time. Real systems proceed in the same way, and here we probably have another intersection with Luhmann: they *entparadoxisieren* (get rid of the paradox) by introducing the dimension of time.

Of course, a system can also get into paradoxes by introducing self-contradictory or mutually negating rules, which it then has to solve. Any programmer has probably had the tiresome experience of realizing that in logically "one-dimensional" computer programs such programming errors lead to a self-blocking or system crash. In programs like our hybrid systems this does not necessarily happen, since they may have the built-in possibility of self-correction. Holland (1995) points to that as well. Here it is true - like it was already described in section 2.5 - that the possible source of self-induced paradoxes can be constructively used just to get rid of these paradoxes again in real time.

A lot more could be said about formal systems in general and hybrid systems in particular. But before this study becomes principally an exercise in general systems theory it is time to turn more explicitly and, above all, more systematically to the possibilities and problems of a social systems theory that can be linked concretely to the concepts, methods, and results that were discussed here.

4. BUILDING BLOCKS OF A MATHEMATICAL SOCIOLOGY

Not only theoretical sociologists, especially in Germany, have a strong aversion to mathematical and natural scientific procedures in spite of the success of the established mathematical sociology, but also the sociological mainstream generally behaves rather abstinently with regard to exact procedures - with exceptions that are yet to be discussed (cf. Fararo 1997).[75] The reasons for this abstinence are partly due to the object itself, namely the already mentioned difference between the systems of the natural sciences and the cultural and cognitive sciences. This was explicitly stressed by Mayntz (1990) with reference to the influence of natural scientific theories on the social sciences; her arguments are basically the following:

The attractiveness of the paradigm of self-organization for systems theoretical sociologists, for example, conceals the fact that a fundamental aspect of social systems gets lost when this concept is transferred to the social sciences: "Only in the social world [...] can a system change purposively the values of its own parameters." (loc. cit., 53) Correspondingly, attempts to transfer the "paradigm of biological evolution" to the phenomena of social change often overlook the fact that people are able "to change social structure and institutions deliberately" (loc. cit., 54). According to Mayntz, that is why natural scientific concepts can only have a limited, and mainly heuristic, value; in particular, the learning ability of individuals and social systems, for which there is no room in natural scientific theories, always has to be taken into consideration.

It should have be clear already that Mayntz' objections largely coincide with what we said about adaptation, learning ability, and the other characteristics of social systems in chapters 1.3. and 2. If the orientation to recent natural scientific concepts only consists in adopting the theoretical and conceptual procedures of the natural sciences, and at the same time neglecting the peculiarities of one's own object, then the whole venture is indeed condemned to end up in purely metaphorical concepts and/or in the reductive analysis of only those phenomena that correspond to the natural scientific "realities". In this context, Mayntz speaks of "isomorphisms" between nature and society; however, only

[75] This can be easily gathered, e.g., from the program of the ISA (International Sociological Association) world congress 1998 in Montreal.

J. Klüver, *The Dynamics and Evolution of Social Systems*
© Springer Science+Business Media Dordrecht 2000

marginal problems can be grasped with these: where sociology is interesting it is not computable, and where it is computable it is of no interest, to vary an old statement.

The analyses concerning the paradigm of universal computability (2.5.) and hybrid systems should have made it clear that in the meantime it has become not only necessary but also possible to extend the scope of formal modelings by the very aspects that are important to Mayntz - and to me. However, this cannot be obtained for free. Primarily, an orientation to natural scientific models can indeed only have a heuristic value, though I personally plead for an adoption of particular *methodical* procedures, i.e., formal models and computer experiments. In the end, as was already mentioned with reference to "self-referential mathematics", the social sciences will not be able to avoid taking their implements from where the theoretical natural sciences have always been taking them, namely from the field of mathematics and logics - in fact, the formal systems. However important stimuli from physics or biology may be, it is equally essential for the social sciences to map their own peculiarities into the most general level of pure systems, and thereby obtain the formal methods that suit them. If this should lead to the fact that mathematicians, logicians, and computer scientists pay more attention to the specific problems of the social sciences and the humanities - selfreferentiality, adaptation, learning ability, self-modeling, etc. - then this could be fruitful for both sides. The history of mathematics shows that this science has always been pushed forward in particular by physical questions; the development of analysis by Leibniz and Newton as well as that of the tensor calculus by Einstein, Ricci, and Levi-Civita are only some striking examples.

This also elucidates why a terminological adoption of single natural scientific concepts can at best be heuristically stimulating, whereas methodically it is almost useless. Natural scientific stimuli often find expression among interdisciplinarily inspired sociologists by the fact that concepts like "dissipative structures", "unbalanced systems", and the like, which describe characteristics of physical-chemical systems, are transferred to social systems, which raises the question of what those concepts can possibly mean in this context. "Unbalance" or "dissipative structures" are known to be defined by the concept of energy in physics. But social systems are not energy processing systems. What should be the energy minimum of a social system? This can only be a metaphor. That is why we have to look for a formal equivalence - for example that of information processing. Correspondingly, Krohn and Küppers (1992), in order to be able to stick to the physical analogies, describe social systems as not being energetically but informationally open.

Such proceeding may be heuristically fruitful, but it still has to be proven whether the analogous use of physical concepts is worthwhile for a precise description of the system, that is, whether an appropriately precise methodology follows from those concepts. However, there are no such proofs yet, and they can hardly be furnished - it is just not possible to map the logic of social systems into that of physical ones. Therefore, I think that it would not make much sense in the long term to look for parallels between the *properties* of physical and social systems. The course I pursue here is a different one, which should have become clear in the meantime: I start out from the most general possible concept, namely the mathematical-logical and, accordingly, that of the social *rule*. Consequently, systems are analyzed with regard to which rule and meta rule systems are characteristic for them. There may be single social systems showing parallels to particular physical-chemical and/or biological ones in their rule-generated dynamics, but such a realization would be the result of the rule analysis and not a *petitio principii*. Accordingly, the premises that were drawn up here about social systems being self-organizing and adaptive can not be sufficiently legitimized merely by reference to historical knowledge; in the end, this legitimization is only possible by showing that self-organizing and adaptive models - computer programs - have a dynamics that is similar to the well-known social dynamics in principle. If one starts out from a mathematical-logical basic concept, namely that of the rule, one has to orient oneself completely to these formal sciences; as a consequence, the transfer of physical concepts into sociology - provided that it is considered to be reasonable - can only mean that in specific cases the dynamics of social systems corresponds to that of particular physical systems, and that it makes sense to use physical models of description - *though always in a merely formal analogy*. However, this has to be proven anew in each single case.[76]

All this could give Luhmann's well-known dictum of "generalization and respecification" (of the concept of autopoiesis) a completely new meaning: the respecification of natural scientific concepts would consist in bringing out the peculiarities of social, psychic, and linguistic systems; and the "generalization" would be the search for formal models that come up to these peculiarities. However, we are only at the beginning of this process. As we could see from the meta parameters, a "self-referential mathematics" requires a new terminology and an extension of the existing techniques. Our concern is not a new mathematics, but one

[76] As a consequence, Bahr and Passerini (1998) are cautious in using the term "analogy" when they introduce concepts like "social temperature" and "social force" which are due to their orientation to statistical mechanics. By the way, this is only possible because their systems are not adaptive.

that deals with the problems described here more explicitly. It is no surprise, for instance, that hybrid systems were given little attention until now: the fact that mathematics and computer scientists that were concerned with dynamic systems mainly concentrated on problems of the natural sciences did not require a systematic analysis of hybrid systems. Only now, with the new sciences of complexity turning more and more to genuine questions of the social sciences, is a shift of attention beginning to show. Therefore, one is permanently forced to work with one's own definitions and concepts. Only concrete research can tell whether this is *sociologically* fruitful in each case. So, let us just try.

4.1. SYSTEMIC STAGNATIONS, REGRESSIONS, AND CONSERVATISMS (TOGETHER WITH JÖRN SCHMIDT)

Up to now, the examples for the dynamics of complex systems from the field of the cultural and cognitive sciences have always suggested something like an unstoppable and constant development of those systems: the continual occurrence of the inconsistency of the relations of production and the forces of production pushes social evolution forward via the mechanism of class struggles; the development of the organism is the continual driving force in cognitive and linguistic development by means of the mechanisms of assimilation and accommodation (Piaget and, accordingly, Chomsky[77]); and biological evolution generates continual development of the species with the same regularity.

Such a suggestion would of course be a much too simple view on the matter. We know many examples from history and from research of the psychology of language and developmental psychology where there was stagnation and even recession instead of a constant development, and the same goes for biology as well (see above, section 3.2.). In particular, such complex systems, which are permanently exposed to developmental pressure, immediately change back into conservative phases as soon as the most urgent problems have been solved. Let us look at these phenomena a bit more closely with examples from sufficiently known social processes.

(a) The constant developmental dynamics of social systems, which were presupposed not only by Marx, obviously cannot be proven empirically. Habermas (1981, vol. I) points out with many examples that neither can the Marxian developmental scheme be maintained empirically, nor does the postulation of a necessarily permanent development apply to all cases. Obviously, there are many examples of social systems that stagnated or even regressed at a particular level of

[77] I do not intend to claim that the positions of Piaget and Chomsky are identical, they "just" seem to be largely similar with regard to this basic logic.

development. Feudal Japan towards the end of the 16th century is a famous example for a "voluntary" stagnation: there the Shogun Tokugawa Ieyasu decided after the first contacts with Portuguese merchants to seal Japan almost completely off from foreign contacts and to freeze the stage of socio-technical development. In fact, he succeeded in the sense that this voluntary stagnation was not ended until the forcible opening of Japan by the USA in the 19th century. Similar developments, though not so radical, seem to have taken place in China after the expulsion of the Jesuits (Spengler 1979). In the language of hybrid systems, Tokugawa set the κ parameter, that is mainly h of course, at zero.[78] Habermas (loc. cit.) gives corresponding examples from different cultural areas.

The cases of social regression, for which we can find classic examples in several American Indian cultures from the South West of the USA (Gumerman/ Gell-Mann 1994) as well as in the Central American Maya cultures, are even more dramatic than the stagnations. It seems that even today the exact reasons for those regressions have not been completely cleared up; however, in the case of the Maya the following scenario seems relatively probable (Gell-Mann 1994):

In spite of their high scientific and intellectual level, the Maya seem to have failed to give much attention to elementary technologies, especially agricultural economy; apparently, they knew neither the wheel nor the plow. As a consequence, their highly developed urban culture had to rely on the technically primitive agricultural work of many dependent peasants. These were forced to inconsiderately overwork, and thereby exhaust, their fields, and so new areas of arable land constantly had to be developed by clearings. Thereby, the ways of transport into the cities became longer and longer, which increased the vulnerability of the whole socio-economic system. In the end, possibly precipitated by climatic changes, the cities could not be sufficiently supplied anymore; their social structure collapsed and the Maya culture regressed to a simple agrarian society, as discovered by the Spanish conquistadores centuries after the regression. The cultures in the South West of the USA may have regressed in a similar way.

With SOZION this development can be simulated quite easily. We assumed that the society's natural environment had become impoverished for the mentioned reasons (or for other ones), so that the function of food production became extremely important for that society; technically speaking, this means that the environmental vector of the component

[78] Admittedly, my knowledge of Japan at this period is mainly taken from the famous novel "Shogun" by James Clavell. However, Japanologists assured me of the novel's great historical authenticity in this respect.

food production was increased, while the other components were decreased. SOZION reacted accordingly: from a society with great social differences there quickly developed a less differentiated society, which in practice became a purely agrarian society again after some runs of the GA. Our formal society's capability of variation - though still rather high - could only bring about that the society was able to develop back into the stage of a society primarily occupied with food production.

How could SOZION or the Maya culture have prevented that regression? Mathematically speaking, both seem to have the problem of the single classes' contributions to the system functions being too rigidly determined; neither SOZION nor the Maya could alter this in any obvious way. Two possibilities present themselves: either a part of the society's intellectual competence is transferred to the invention of new technologies of agriculture and transport, so that both the production and the transport of food can be optimized; this would mean a diversification of the intellectual classes' contributions into the system function "food production." Or as many classes as possible put parts of their contribution value B_{ij} directly into the system function "food production"; this would enlarge the share of productive rural workers, especially for the short-term, and by that increase the output of production as well as decrease the number of unproductive consumers.[79] A corresponding possibility of varying B_{ij} through the GA of SOZION also produced the result that the society largely maintains its social differentiation under these circumstances. However, this second solution is rather unrealistic, since the urban dwellers would have worked in the fields only under duress, and probably not very productively anyway. Such experiments have most recently failed in the Maoist Cultural Revolution and in the "primeval communism" of the Khmer Rouge in Cambodia.

In formal systems like SOZION, stagnations and regressions can be easily simulated; in particular, the respective causes can be reconstructed and varied experimentally. Naturally, systems like SOZION are still extremely simple; however, they can make the common fundamental logic in the dissimilarity of different system developments become transparent. In the case of regression analyzed here, the common logic apparently consists in the fact that in both systems - SOZION and the Maya society - the functional position of the single classes could not be changed, which can be seen from the constancy of the B_{ij}. If the B_{ij} would have been kept variable, then obviously other possible solutions could have been found. Thus, in the case we are dealing with here

[79] Let us recall the evaluation function of SOZION here: the system value W ensues according to the formula $W = (\Sigma A_l * B_{lj} * F_j, \ldots, \Sigma A_k * Bk_j * F_j)$, where B describes the contributions of the single classes A to the system functions F.

particular constraints, namely the social immobility of a society, prevented the realization of mathematically possible solutions. For an explanation of these constraints we would probably have to make reference to human interests, namely the privileged urban dwellers' aversion to take the comparatively arduous life in the country upon themselves. However, it applies once again that the *mathematically* most simple solution is certainly the one that SOZION has realized with a constant B_{ij}: the system did not need to change very much.

A much more interesting object of study is a social - and probably not only social - phenomenon that seems to be universal, and that can be described as the immanent conservatism of societies. What that means is the following:

When a social system has reached a favorable state by variation of its rules - favorable with regard to the requirements of its external and internal environment - then the problem arises of how this or another favorable state can possibly be maintained. It is in fact possible that a rule variation - in the case of high values of κ and/or μ - causes the system to attain a highly complex dynamics; if the rule variation is stopped now, the very rules that caused the attainment of the favorable state can subsequently generate a trajectory that quickly leads away from the favorable state. We saw that this is always the case, for example, when a control parameter like P is changed so that a favorable state is reached which "normally" could have been realized only by a rule system with a low P value. In this case, the rules must be varied anew, which can again lead to a high complexity of the system dynamics, and so on. As a result, the system would never come to the point where it could generate its sequence of states with constant rules for some time, but would permanently have to undergo rule variations. However, rule variations in all systems - even in the computer - can only be performed at the expense of an increased expenditure of energy (i.e. more computer time), and therefore such a system would probably soon be hopelessly overcharged. The history of rule variations, especially that of rather radical ones (that is, revolutions) really shows that periods of relative rule constancy returned quite soon after comparatively short phases of change - one just had to recover first. The same can be observed in connection with the genesis of new types of societies like, e.g., the European feudal society: after it had once emerged, no effort was spared to change its rules as little as possible.

There are of course different solutions to this problem; we shall call a particularly simple one an "elementary solution": if a favorable state is reached by rule variation, then practically all important rules are switched off, so that the trajectory deviates from the favorable state only

insignificantly - the latter is practically forced to become an attractor of a low period. This elementary solution can be observed again and again throughout history, e.g., in connection with the problem of social differentiation: in the developmental stage of differentiated societies it is quite easy for particular individuals to realize social advancement - and sometimes also deterioration. However, the more differentiated a society gets, the more rigid are the regulations for transitions from one class to another, and in caste or estate societies such transitions are almost made impossible. The transition matrix, i.e., the matrix that contains the probability values for social mobility almost exclusively comprises elements with values close to zero, to speak in terms of formal systems such as SOZION.

One might of course explain the social systems' preference for the elementary solution with human character traits like active property preservation, striving for power, competitiveness, orientation to the family, etc, which would describe the point reasonably well. However, there is also a mathematical explanation for the fact that there is a certain inevitability attached to the elementary solution; this can be explained quite well with the system SOZGEN:

We learned that SOZGEN has the option of switching the GA off or on, depending on how favorable its states - or more precisely: the states of the CA - are. Furthermore, the biological basis of any society is taken into account in SOZGEN, that is, the fact that individuals die and new individuals are born. In our experiments with SOZGEN we could observe the following: when the CA reaches a favorable state and the GA is switched off the CA usually has a trajectory in which the new initial state, i.e., the state at which the GA was switched off, changes only slightly. By that, SOZGEN is able to reach rather long trajectories of the CA without the GA, that is, a long sequence of favorable states - SOZGEN constantly manages to pause the coming rule variations, as it were. An analysis of the relevant rules in these "quiescent phases" shows the following: those trajectories are realized by the fact that either most rules are switched off, or almost all probability parameters are very close to zero, so that there are only a few interactions between the cells.

At first glance, this result is surprising, because the GA of SOZGEN did not receive this as an optimization goal. The GA exclusively orientates to the same evaluation function as SOZION, which does not include any favorable *trajectories,* but only favorable single states. However, if we recall the functioning of the GA, and particularly the "contractive" proof by Michalewicz (3.2.), the riddle is solved quite quickly:

When the GA is told to optimize rather unfavorable states at the beginning, then $d = |U - V|$ is relatively great. In order quickly to minimize

d, radical variations are required, because a new set of rules is always tested with at least three runs of the CA, even those sets of rules that cause the most radical changes of states. On the other hand, the closer SOZGEN gets to favorable states, the smaller the variations have to be and, in particular, the rules of the CA must minimize the latter's changes of states as well. If this would not be the case, the respective CA would quickly deviate from the favorable state again during the test phase, so that the corresponding set of rules would have to be sorted out. Thus, the GA optimizes the CA not only with regard to small d values, i.e. favorable states, but also with regard to rules that cause the slightest possible changes of states; one could say that it contracts both the favorable solutions and the sequences of states themselves by permanently reducing the effect of the rules. This result was actually not intended, and so we were in fact surprised that the GA gave us the explanation for the preference for the elementary solution as well - "getting something out of nothing", to quote John Holland once again.

The reason why the GA behaves this way, and the fact that this is not only due to its specific way of functioning, can both be beautifully illustrated with a classic argument which was used by R. A. Fisher, one of the founders of biomathematics, to refute so-called "saltationism" in evolutionary biology (cf. Dawkins 1987); the saltationists try to explain biological evolution by the assumption of "macromutations", that is, jumps of nature as it were. Fisher used the "microscope argument" to prove that this is mathematically impossible: if someone wants to sharpen the focus of a microscope that is already focussed fairly well, he may make only slight changes, because otherwise the probability of an improvement would decrease immediately. If, e.g., optimum is at about 0.5 and the present state is 0.45, then a change of 0.2 will always render a worse result - 0.65 or 0.25, depending on the direction of the change - than a change of 0.05 - 0.5 or 0.4. Therefore, it is a mathematical necessity for nature to favor micromutations in order to achieve good results in adequate time.[80]

[80] Dawkins gives other arguments against the saltationists as well, and Fisher's argument is in fact more subtle than it was presented here. However, biological subtleties are not the point here. By the way: Fisher's beautiful example could actually not prove the untenability of the position of the saltationists, which obviously did not occur to Dawkins: the adjustment of a microscope or any other corresponding instrument is a variation of a *linear* and deterministic system, whereas biological evolution is evidently neither of the two. Only by proving that the same goes for non-linear and stochastic systems could the saltationists be refuted. In this respect, we can take the results regarding variable meta parameters in section 3.3. as a late justification of Fisher's arguments, since all our results were gathered from non-linear and partly stochastic systems.

Fisher's argument can be understood as saying that nature has to proceed contractively, because optimization processes generally require that. In this respect, the learning rules of neural nets are also contractive, as is any optimization algorithm that orientates its proceedings to the value of the previously attained solution. Hence, the GA in SOZGEN always supplies the elementary solution because the hybrid system's meta operators must pay attention to favorable rules *and* because the program can compute the value of the rules only from the states of the CA. According to Fisher's microscope argument, this results in the fact that rules which cause favorable states plus radical changes of states cannot be as good - in terms of probability theory - as rules which produce favorable states plus minor changes of states. The former just have a too great probability of the states becoming worse again in the test phase.

This is obviously another example of the variation of meta parameters that was analyzed in general in section 3.3.3., namely a variation in accordance with the delta rule in neural nets. The phenomenon of the elementary solution observed in SOZGEN can be understood as a radicalization of the r parameter, since SOZGEN is evidently not only concerned with taking more and more rules away from variation, but even to block more and more rules. Therefore, it seems reasonable to introduce an additional meta parameter b here for the blocking of rules: if we call the number of rules that are in force in a base system at a specific time k_1, and the corresponding number at a second time k_2, then for this time interval we have

$$b = k_1 - k_2.$$

It is clear that b is another meta parameter of the "variable" type, which like the variability of μ has been established by the system "itself".

Thus, according to Kauffman's "Kantian stance" we can say that social systems favor the elementary solution because it is the mathematically most simple one. Consequentially, social systems become rather conservative and rigid after rule changes, because otherwise they would have to keep on responding with new rule variations (see the coercions to reforms in democratic societies in section 3.1.2); the mathematically favored elementary solution, which is not at all trivial in light of the above considerations, practically forces the systems to behave that way, which makes it unnecessary to fall back upon human character traits for

an explanation.[81] Of course, it cannot and shall not be denied that social interests always play a part in the consolidation of particular societies. However, we can at least say that hybrid social systems - *provided they are not stopped* - always tend to the conservative establishment of their currently realized conditions - regardless of whether the protagonists are actually conscious of that or not. Hence, it is a political question whether one wants to accept this mathematically obvious solution or not. Thus, a sociology that attempts to pursue new paths in the tradition of the mathematical and computerized sciences should see one of its most important tasks in pointing to these problems: it is not completely wrapped up in that tradition.

When the dynamics of SOZGEN is observed for a longer period it turns out that the state values gradually worsen so that sooner or later SOZGEN has to switch on the GA again. In fact, not all the rules have been switched off, so that a non-trivial dynamics - not a point attractor - remains. The reason for this is that the test phases of the CA without the GA are relatively short; prolonging them considerably would lead to a total standstill in the long run. The present kind of operation is certainly more in keeping with social reality, since completely stagnating societies hardly exist. It may be true that simple tribal societies behave that way, but the societies that we are dealing with here have already gone through serious changes by rule variations; these societies always show some changes, that is, non-trivial dynamics.

SOZGEN shows that these changes of states always lead to a worsening of the system values in the medium- or long-term, which sooner or later requires the use of rule variations again. Mathematically speaking, this long-term failure of a "carry-on-policy" results because there simply are many more bad possibilities than good ones. SOZGEN may well improve with constant rules for the short-term; in the medium- and long-term, however, most systems get caught up by the law of probability, so that *even in a constant environment* the conditions either become worse or the system takes the strain of rule variations - reforms or even revolutions - upon itself. It is obvious that the rule variations have to be more radical the longer the system delays the changes, and the

[81] In connection with experiments with CAs Schelling (1971) likewise pointed to the fact that it is not necessary to make assumptions on racial discrimination etc. to explain the formation of ghettos and the segregation of ethnic minorities respectively. One simply needs to assume that people prefer living together with people who are similar to them to living with people who are not; on that condition we can learn from simple CAs that this automatically leads to segregation. Of course, there is still the question of where the criteria for similarity come from; they certainly are socially determined and variable themselves.

worse the system values have become as a result. It should be obvious that this is true a fortiori for variable environments.

It was already mentioned that in contrast to SOZION, SOZGEN has a "biological basis", i.e., an additional dynamics of population. SOZGEN can vary the birth and mortality rates of its population according to their specific class; our observations showed, however, that such variations do not notably affect the dynamics as a whole. But the fact that such a dynamics has been taken into account really does have an effect: SOZGEN's dynamics is much more complex than that of SOZION. If only because of that there is the necessity of always taking aspects of the dynamics of population into consideration when analyzing the dynamics of social systems; unfortunately, this is seldom the case in theoretical sociology. The distribution conflicts between the generations (pensions vs. training vacancies) that can be observed in Western industrial societies point all too obviously to this deficiency.

There is another important aspect concerning the above subject of social elementary solutions: actually, social systems are only able to realize the elementary solution because they have an additional rule which allows them to realize favorable states via generational succession as well. This concerns the fact that the social position of parents is passed on to their children in principle, that is, with a high probability. This comes out most clearly in caste societies like in classical India; less rigid, but still very effective, it can be observed in feudal estate societies, and even "open" societies (Popper) like modern Western societies are marked by such rules. It is perfectly clear that a society can switch the other rules of social mobility off, or make them less effective, only by means of such a "traditional or hereditary rule": if such a rule did not exist, then rules of social mobility would have to be switched on again, varied, or newly introduced at every alternation of generations - and that means continuously, in real societies. Thus, traditional rules, which determine, of course, whether an individual has access to or is excluded from social privileges, ensure the system's conservatism over the generations, and, if only for logical reasons, systems can do without them only at the expense of the impossibility of rule constancy. Since practically all societies operate with rules that favor the children of privileged parents, and vice versa, the question arises again whether these universal patterns are biologically founded - parents try to keep their children in social proximity because that is how it is done in all biological communities - or if this shows another mathematically necessary feature of dynamic systems: if a society realizes the elementary solution in the above sense in order to undergo as few changes as possible, then the introduction of such a traditional rule is a necessary result. This would be an illustrative

example of a social universal in the sense of the Chomskyan "formal universals" (Chomsky 1965); the above considerations suggest that.

SOZGEN usually works with such a traditional rule, which can not be varied by the GA. In particular, two individuals of opposite sexes can only have children in SOZGEN if they belong to the same class, or to classes that are not too different from each other - surely a realistic assumption. For the children, this means that they obtain the status of the higher class if their parents belong to different classes; otherwise they inherit the parents' class status. Since we wanted to find an answer to the question of the *logical* necessity of the traditional rule, we kept the "marriage rule" constant in SOZGEN, but left it to the GA to determine to which social class the children belonged. Once again, it has to be noted that SOZGEN is a stochastic system; the results of these experiments can be generally summarized like this:

SOZGEN really realized the simple traditional rule comparatively often, which means that it put the descendants into the class of one of the parents (in this respect, SOZGEN makes no distinction between women and men). However, this can not be observed as often as the mentioned blocking of rules; furthermore, SOZGEN is forced to switch on the GA again rather soon, that is, to make new social adaptations as it were. We assume that SOZGEN would realize the traditional rule more often in longer test phases (see above); this is still to be tested though. But since traditional social systems usually do not take much time to test such rules, it can really be assumed - just because SOZGEN does not permanently realize the well-known social hereditary mechanisms - that the well-known hereditary rules from traditional societies have to be put down to biological constraints. It is of course another question how the corresponding facts can be explained in modern societies.

4.2. SELFREFERENTIALITY AS SELF-MODELING (TOGETHER WITH CHRISTINA STOICA AND JÖRN SCHMIDT)

In section 1.3. it was stressed that, and to what extent, the capability of self-modeling is a fundamental aspect of the selfreferentiality of social systems. Of course, the term "model" is not used in the conventional sense of model theory analysis here, that is, as a goal-directed and planned design. Informal self conceptions of societies as "the land of the free and the home of the brave", "la grande nation" or "Deutschland, heilig Vaterland" are models to which a society orientates its developmental dynamics and, most of all, its own evaluation functions. In particular, the decisive factors for changing social rules or keeping them constant, which were brought up in 3.3.3., are probably

fundamentally connected with these self conceptions. The Balinese institution of the cockfight mentioned in 1.3. can be interpreted as a self-model in which the Balinese society becomes conscious of itself, confirms itself, and thus protects the adequacy of its own rules.

"Reproduction theoretical" analyses of society have often given more attention to this problem, though in a different terminology of course. This can be illustrated with Bourdieu's famous analyses of the social systems of education and science (Bourdieu/Passeron 1964; Bourdieu 1976):

According to Bourdieu, the main social task of institutionalized education in general, and scientific education in particular, is to ensure the reproduction of the most important social structures. The term "reproduction" implies two meanings here: on the one hand, it concerns the fact of keeping the basic structures of society in general stable over the generations by means of the subsystem education; this is mainly achieved by the development of specific habituated action patterns which pre-form socialized individuals, so that they constantly renew the established social rules, that is to say, they behave according to a constant and traditional rule system. This was also brought out by Parsons with regard to the L-subsystem, as was mentioned in 2.4. On the other hand, the basic structures of society must themselves be realized in the educational subsystem, in order to guarantee the effectiveness of the habituated action patterns. Hence, the social system of education performs its social task of reproduction by the fact that as a subsystem it is a model of the society as a whole, and thus the society models itself in the educational system.

Correspondingly, the subsystem science can be understood in terms of reproduction theory: the differentiation of the social system of science into distinct disciplines and complexes of disciplines, respectively, represents the system's reaction to the social differentiation into distinct subsystems. Obviously, the natural and engineering sciences correspond to the subsystem economy; jurisprudence corresponds to the legal system, and so on. The single disciplines' "*Leistungen*" (achievements), to use Luhmann's term, ensure the long-term reproduction of the particular subsystem they are assigned to, in fact in the specific fashion of their discipline - innovations in the economic system, legitimations in the legal system, etc. At the same time, the scientific disciplines are a model of the different social subsystems in the sense that the latter's basic social structures - action rules - are repeated in the social organization of the disciplines. With that, the concept of science as social acting on trial, which was mentioned in 1.3., gains a precise sociological meaning.

It can be concluded from these reproduction theory outlines that self-modeling can be found as a basic principle in various social formations,

and that it seems to be a fundamental part of complex social systems. It is certainly no accident that institutionalized cases of self-modeling often have the function of testing something without having to risk a societal emergency. One might of course object that especially in modern societies, which are under continual pressure of competition, a failure of education and research must have disastrous consequences. But in the areas in question, a short-term developmental failure does not directly make its mark on the whole society. It is relatively easy to take remedial action here - provided that the society as a whole summons up its will, which can of course not be taken for granted.

However, rather than following these reproduction theoretical thoughts in our experiments of simulating the self-modelings of social systems, we were governed by assumptions of model theory, as they are common in the mathematical sciences, and especially by the learning theory analyses of Bateson (see above, section 2.3.). In accord with Bateson's learning III, we were particularly interested in analyzing the effects that ensue when a specific program transforms itself into another program and thus takes the second program as a model of itself. This is not exactly learning III as regards content, since that implies an irreversible transformation, whereas in our simulated self-modeling the first program comes back into effect after having been modeled by another program. The heuristic force behind that was Bateson; however, it seems to be possible to formally grasp the principle of personality transformation through learning III, which he mentions, by self-modeling programs.

The program that was designed to model itself is the meanwhile sufficiently explained hybrid system SOZION. This system was horizontally coupled with another hybrid system, namely an interactive network (IN), and vertically coupled with a GA. So, on the whole we have a horizontally-vertically coupled hybrid system, and thus a second order hybrid system whose parts are hybrid systems themselves. As was explained in Klüver 1995, interactive networks are neural nets which are feedback networks and whose weights, unlike those of learning networks, do not change. Their activation rules are the linear standard rules described in section 2.3.:

$$A_i = \Sigma A_j w_{ij} \qquad (4.4.1.)$$

The states of the cells in the CA are understood as symbolizations of the particular individuals' social roles here. A self-modeling of SOZION's CA that does not simply repeat its bottom-up approach would be the construction of a model that only includes the CA's macro-level, and therefore does not represent the roles of the single individuals, but

only the different roles' aggregations. Since these aggregations have been described as social classes, we get a network as the model of the CA whose elements are the classes of the CA. The changes of the CA's states are then modeled as changes of the network's states according to the formula (4.4.1.). Hence, SOZION models itself by mapping its actual states into a network, with the size of the single classes corresponding to particular activation values of the corresponding network elements; the probabilities of upward or downward mobility between the classes correspond to particular weights in the net (w_{ij}); and the changes of states determined by the CA rules correspond to the respective changes in the net. The following is a more precise description of the process of self-modeling:

As usual, the CA of SOZION is changed by the GA at the start of the whole program with regard to an environmental vector U. After a few runs of SOZION, i.e. CA with GA, the CA transforms itself into an interactive network (IN). The net consists of ten units because SOZION can develop a maximum of ten classes. At the beginning of the transformation the program computes the number of members of the individual classes, and puts these numbers in proportion to the total number of cells (4000 on the whole). These proportional values make up the external activation vector for the IN[82]. We already mentioned that the weight matrix is the same to the IN as the matrix of the transition probabilities in the CA.

With that the transformation CA The transformation CA $\rightarrow$ IN is thereby already carried out. The IN corresponds to the CA that had the best system value after the optimization as its macro-model; the external activation vector corresponds to the (standardized) final state of that CA before the transformation. Now the resulting IN is also optimized by a GA with respect to the same environmental vector U. This happens somewhat differently than the optimization of the CA: the weight matrix is written as a vector on which the GA *directly* operates; thus, the vector's components are not ternary coded, but instead are the real numbers of the w_{ij} (rounded to two places). For the operation of the CA this means that the variations through crossover can happen much more radically - as was mentioned in section 3.3.3., s is much greater in a hybrid IN than in SOZION. So, things are much rougher in the model than in the "reality" of the CA. Subsequently, the IN optimizes itself

[82] Let us recall in what way interactive nets operate: at the start, all units have the value zero. The start of an IN is effected by the "external activation" of single or all units, which means that they are given a specific initial value that is not zero. The IN runs through different phases according to the activation rules (4.4.1.) until it reaches an attractor state of the period zero - it "stabilizes". What this means is of course that all units reach activation values that do not or only insignificantly change in further runs.

analogously to the optimization of the CA and, after some optimization steps, transforms itself back into the CA - in fact, into the state from which the CA had transformed itself into the IN. It adopts the IN's optimized weight matrix as its transition matrix, optimizes it with the GA, transforms itself again after some runs into a new IN that optimizes itself accordingly, and so forth, until a goal state is reached, or the optimizations no longer improve considerably.

When we constructed the whole program, which was called SOZAIN - SOZION as a CA plus IN - we really did not know which optimization results the program would achieve. We did not actually plan on any particular effects, because the rules of CAs and INs are very dissimilar, and therefore it could not be expected that the two subprograms would yield any useful results to each other. The optimization results of SOZAIN are all the more remarkable when compared with those of SOZION:

SOZION usually starts with differences of d = 0.5 to the environmental vector U, and optimizes d after a total of about 80 runs of the CA (including optimization steps) to d = 0.08. Subsequently, there are no considerable improvements. The reason for that is that the vectors of the GA have basically become similar, so that the crossover operator cannot produce any effective changes of the rules. If the mutation rate is increased, little improvements can be achieved, but they are not significant, and they also require longer runs of the program.

The optimizations in SOZAIN take a quite different course. This program already achieves values of d = 0.01 and better after significantly fewer than 80 runs; values of d = 0.002 and under are not unusual if SOZAIN is allowed to run a little longer. In SOZAIN the values of d continually improve with longer running periods; the effect of longer running periods not achieving better results, which was observed with SOZION, either did not appear in SOZAIN, or occurred much later. We have not yet found an absolute optimization limit for SOZAIN; when the outcome is rounded to three places, values of d = 0.000 are not unusual. Such good optimization results even occur when the IN itself does not improve considerably; since the d value for SOZAIN is only measured with respect to the "real" CA, it does not seem to make any difference if the IN realizes favorable matrices "only" for the CA, and not for itself.

As we said before, we did not reckon on such results. The differences between CA and IN did not give rise to the assumption that such a good correspondence between the two programs could be established. On the other hand, the two systems are "dynamically similar", because they both belong to complexity class II (the IN does not always reach absolute point attractors either). This has already been mentioned with regard to

the CAs of SOZION and SOZAIN respectively. Thus, we had to check whether the results could be explained by the dynamical similarity between CA and IN. With regard to that, we changed SOZAIN so that, as a result of minor rule changes, the CA belonged to class II *at every initial state*, and repeated our experiments. However, the results were completely identical to those of the original SOZAIN, which shows that dynamical similarities of the base systems are unimportant here.

Undoubtedly, the explanation for our results rather lies in the fact that we are dealing with a meta parameter here, in fact, with the s parameter that was already described in section 3.3.3. SOZION's relatively high s value guarantees comparatively good results, which is different from the hybrid IN: the latter, with its low s value, even tends to worsen regularly after being quite successful at the beginning, so that the IN in itself is clearly inferior to SOZION with regard to its adaptive success; the fact that the IN is actually the simpler of the two systems regarding quantity makes this all the more remarkable. The results can be interpreted to say that a very low s value is unfavorable because the serious variations constantly make the system deviate from already attained favorable states. This clearly shows that even an optimization algorithm as effective as the GA has its limits.

On the other hand, the high s value of SOZION is not suited to cross a limit once it has been reached - in a certain sense SOZION is too irresolute. The greater robustness of the hybrid IN is useful here, since it helps SOZION to find its way out of evolutionary blind alleys. For SOZION, the optimized matrix that it gets from the IN, and that is usually not particularly favorable for the IN, is the way out which it cannot find on its own with its cautious procedure. Thus, s shows itself as a variable parameter, though it does not necessarily follow the adjusting principle of the delta rule, that of μ or b (see above, section 4.1.), but has to be varied, it seems, whenever there is no possibility to proceed in the cautious way. As a result, the variation of s corresponds to the well-known necessity of leaving an established and familiar procedure in case of emergency and trying to go on in a completely different way - in this respect it is comparable to Bateson's learning II or even III. In a very formal parallel, the program SOZAIN can also be understood in the sense that the hybrid CA "consciously" models itself in order to tentatively carry out the changes of s which it would not be able to do in its original version.

Of course, it would have been possible to carry out the variations of s in SOZION itself without taking the roundabout path via the IN. However, the logic of SOZION merely consists in a high s value, and we thought it quite interesting to study the effects that a simplified model of a formal system has on that system. The results are, at least, not

unsatisfactory. By the way, we should remember (see above, section 3.3.3.) that SOZION alone already achieved equally good optimization values by a variation of the g parameter. It seems that the two parameters have similar effects.

Furthermore, the results regarding SOZAIN can be taken as classic examples of the concept of emergence described in section 2.7.: the two hybrid subsystems of SOZAIN were perfectly well-known to us, and have been examined in numerous experiments. The coupling of the hybrid CA with the hybrid IN is nothing special either, since it is just a transformation of equal and standardized values from one subsystem into the other and back. However, despite of our knowledge of all the relevant parts of SOZAIN, we could not completely understand its total behavior after the first results, so that we could only offer more or less well founded assumptions at first.[83] Thus, according to the above definition, the behavior of SOZAIN was emergent *to us*, because it could not be predicted or explained by any explicit theory. The explanation, which is relatively simple *in retrospect*, only resulted from our systematic occupation with the different meta parameters. We assume that this is quite common in connection with the explanation of emergent phenomena: in retrospect they are very comprehensible.

So, on the one hand, the *total behavior* of SOZAIN clearly demonstrates why it only makes sense to speak of emergence with regard to (as yet) non-existent explanatory theories; however, we had no doubt that SOZAIN could be explained in the end. On the other hand, in contrast to the above mentioned position of Cariani (section 2.6.), one can see that computer programs can in fact show emergence in that sense. They convey new insights regarding the behavior of complex systems, which could not be predicted as such, and which particularly gain their new quality from the difference between the parts of a system and its total behavior. However, the very fact that there was hardly any doubt about the total behavior of SOZAIN being mathematically explicable shows once again that there is nothing mysterious about emergence: if emergence occurs in this sense, then this is merely an indication of an as yet insufficiently complete pattern of explanation.

Basically, SOZAIN is still a rather simple system, and the fact that emergent phenomena even occur in systems that are actually quite transparent might point to a more fundamental problem: compared to actual complex systems such as societies or the brain/mind, the systems that we experiment with are still so simplified that it is certain that

[83] These are described in Klüver/Stoica/Schmidt 1997; because of an imminent publication deadline we were forced to publish SOZAIN before we could gather any systematic results.

surprising effects similar to those of SOZAIN will regularly occur in more complex programs. Therefore, it is quite possible that computer programs which are only nearly as complex as their real models will be as incomprehensible as the real complex systems presently are. Such a possibility has already been considered by Stadler and Kruse (1992); they noted that if there really were a factual artificial intelligence, then its designers would be unable to understand it, just like we are presently unable to understand "real" intelligence.

With that, Stadler and Kruse argue against the essence and purpose of AI research and, above all, against the systematic construction of highly complex computer programs for the analysis of real systems. However, the experiments with SOZAIN allow yet another conclusion, namely that simplified models can in fact be relevant for their richer originals. The IN is a reduced model of the CA inasmuch as the entire geometry of the CA is not taken into account in the IN; its relevance has been frequently emphasized, and we will come back to it in more detail under the heading of the geometry of spaces of social interactions. Furthermore, the IN's GA cannot switch off complete rules like in the hybrid CA; it can only annul single rules by setting specific w_{ij} at zero. Despite these essential simplifications, it is evidently possible to get relevant values for the CA from the IN. Therefore, SOZAIN in fact encourages a deliberate use of simplified models that only bear a faint resemblance to their original. As long as one does not *equate* reductive models with their originals, which unfortunately is a widely held attitude in AI circles, SOZAIN shows that simplified models are quite reasonable. However, the frequently noted research assignment of a systematic investigation of those connections between models and originals remains an important task.

Let us finally return to the sociological dimension of self-modeling. The dynamics of a CA are "bottom-up" - a society is modeled by the interactions of its elements on the local level. The IN of SOZAIN, by contrast, is a "top-down" model: only its global aggregations are observed, and the consequences for the local interactions are derived from their mutual relations. That is why SOZAIN can equally well be interpreted politically as a model for a policy "from below" and "from above", respectively: the IN plays the part of a central planning department as it were, which lays down rules of local interactions on the basis of global analyses. The results of SOZAIN show that such an undertaking can really work, especially if (and, in fact, because) basic local aspects have not been taken into account in the global model.

The history of global social planning is in principle rather negative, since it is connected with dark historical phases such as, e.g., the totalitarian regimes of the 20th century. On the other hand, "soft" forms

of social planning, such as, e.g., Roosevelt's New Deal, show that they can also be successful and, in particular, can take place without bloodshed. It is certainly not possible to carry out real planning with a program such as SOZAIN; it is constructed much too simply for that. But perhaps it is possible to discover the conditions of the possibility, to recall Kant once again, of really systematic and democratic planning. At least SOZAIN is the beginning of a possible way toward such a goal.

4.3. GAME STRATEGIES AND SYSTEM DYNAMICS: SOME THOUGHTS ON THE RELATIONS OF THE THEORY OF GAMES AND SYSTEMS THEORY (TOGETHER WITH RALPH KIER)

Despite the research tradition of mathematical sociology that was mentioned in the introduction, and in spite of empirical social research, with its established use of statistics and appropriate computer programs, there is, in general, no doubt that the frequently mentioned deficiency of *general* sociology regarding formal methods really exists. There are of course exceptions, and two of the most important ones must be described here: first, there is the (mathematical) theory of games developed by John von Neumann and Oscar Morgenstern; and second, there is the analysis of social networks. Both research tendencies can look back on years of tradition in the use of formal techniques and, consequently, both have richly developed formalisms. We shall be dealing with the analysis of social networks under the heading of the "geometry of social spaces;" the subject of this subchapter is the fruitfulness of game theoretical approaches for a theory of social dynamic systems.

At first glance, it seems that game theoretical analyses do not have much in common with our subject, and it is certainly no accident that sociological systems theorists have only rarely taken notice of the game theoretical approach. Therefore, it is by no means a matter of course that, e.g., Rapoport in his "General Systems Theory" (1986) actually names the theory of games as the only area of systems theory application in the social sciences, without explicitly substantiating this claim. The classical version of the theory of games is not concerned with the dynamics of complex systems, but instead with the rational and principally egocentric behavior of single actors, with no consideration for systemic contexts. Correspondingly, social scientists who use game theoretical approaches are usually explicitly or implicitly under an obligation to action theory traditions in accordance with the rational choice approach. Terms and concepts of systems theory just do not occur there.

However, the connections between systems theory and the theory of games became very clear primarily in the works of Axelrod (1984 and

1987) and Maynard Smith (1982); in the meantime it has almost become taken for granted that game theoretical analyses employ a systems theoretical terminology, and that game strategies are defined as parts of complex systems and generators of equally complex system dynamics. Leinfellner (1995) even characterizes the modern "evolutionary" theory of games as the actual basis for the analysis of social dynamics. Let us therefore take a closer look at the basic concepts of the theory of games:

A "game" in the theoretical formal sense is usually defined as a situation of decision where two or more actors have to act against or with each other. An analysis of this situation of decision, according to the classical approach of the theory of games, consists in discovering the particular behavior for each actor that is the most advantageous *under the given circumstances*. The different possibilities of behavior are usually described as *strategies*. The basic assumption in game theoretical analysis is generally that the actors act "rationally" and "egocentrically", that is, they orient their strategies to the greatest possible and recognizable advantage, with no consideration for other aspects of the situation. Thus, the point is just a weighing up of goals and means; the famous definition of strategic-instrumental action by Habermas (1981, II) applies perfectly here.

In order to make such rational behavior possible it is necessary that the pros and cons of single strategies can be recognized and calculated against each other. In section 3.2. we already described the iterated prisoner's dilemma as one of *the* classic examples of such games, that is, situations of decision; there the pros and cons of cooperative or aggressive behavior were described in the form of a so-called "pay-off matrix." Let us look at that in a generalized form once again for clarification:

	C	*D*
C	*a*	*b*
D	*c*	*d*

with C representing a cooperative player and D an aggressive one.

In general, we can conclude from the above assumption of the possibility of defining and recognizing the pros and cons that a game can always be described by such a pay-off matrix (though other formal descriptions are of course possible and common as well): there are particular basic types of strategies, cooperative and aggressive in our case, and for every move in the form of a basic type it has to be defined what its result is with respect to a "strategic interaction" with the player who uses the same or a different basic strategy.

According to Axelrod (1984), we can describe any game of the type of the iterated prisoner's dilemma by the fact that it is true that

$$c > a > d > b \qquad (4.5.1.)$$

Most games of this type are analyzed according to this condition; in section 3.2., e.g., we had $c = 5$, $a = 3$, $d = 1$, and $b = 0$. This is not a necessary definition though. Miscellaneous analyses of the iterated prisoner's dilemma vary the condition (4.5.1.), as we shall see below.

If we examine the basic concepts, it is quite clear why the theory of games traditionally has belonged to the context of decision theories, and not to systems theories. However, this is not necessarily so, and the application of the theory of games to evolutionary biology by Maynard Smith already shows that the theory of games can be used in a much more general sense:

> "In this chapter I introduce the concept of an 'evolutionary stable strategy', or ESS. [...] The concept is couched in terms of a 'strategy' because it arose in the context of animal behaviour. The idea, however, can be applied equally well to any kind of phenotypic variation, and the word strategy could be replaced by the word phenotype; for example, a strategy could be the growth form of a plant, or the age at first reproduction, or the relative numbers of sons and daughters produced by a parent." (1982, 10)

If we generalize this terminology beyond its biological context, then the game theoretical concept of strategy - and with that its connected formal determinations - can be applied to any form of *interaction between a unit and its actual environment*. In fact, the biological meaning of "phenotypic variation" is just the adaptive interaction of an organism or a species with its actual environment, and also, in particular, the specific forms of interaction between individuals within the same species. As a result, we already have a conceptual relation between systems theory and the theory of games through understanding strategies as specific interactions between system elements, or systems and their environment.

In previous formal modelings of social systems, strategic interactions between the elements have not been taken into account. The basic idea was that a social system had to form rules of interaction between its elements on the basis of particular environmental requirements in order to achieve these requirements. These rules were interpreted, e.g., as rules of social differentiation, and thus as rules for the development and establishment of specific social classes and the roles which go with them. The decision of *whether* these rules are successful for the system depends on the total system behavior; the "role occupants" had no choice but to adapt to the rules and their variations, i.e., to behave according to the rules. Therefore, these models *methodically* represented a bottom-up approach because they were constructed on the basis of local rules of

interaction; however, *theoretically* they follow a constitution "from above," according to Luhmann and Marx: not only does the system determine what its elements are, but their behavior results exclusively from the system rules as a whole.

Such an approach can of course be applied beyond the confines of processes of social differentiation. Lansing and Kremer (1994) analyzed the development of social networks on the island of Bali that existed between the individual villages in order to coordinate the irrigation economy; they reconstructed the basic logic of this development in a computer model that followed Kauffman's Boolean nets (see above, section 3.1.2.). They found out that there was no central planning logic in the development of those networks, but that the single villages acted both "egocentrically", i.e., on the basis of their own interests, and in agreement with their respective neighbors. The functioning of the whole network, especially on variable ecological conditions, results as a system property, inasmuch as it cannot be ascribed to the intentions of the single peasants or villages: "The ability of the system as a whole to react to changes is thus a property of the network itself, rather than the accidental consequences of the actions of individual subaks or temples." (Lansing/Kremer 1994, 217). Again, whether the individuals' behavior is successful or not depends on how well the system functions as a whole, and not primarily on the welfare of the single actors, even if this is naturally their primary orientation. Thus, the development of those networks can also be understood in a game theoretical way as the product of the actors' "social intelligence", which is what Leinfellner (loc. cit.) suggests.[84]

Introducing the concept of strategy as an additional form of interaction obviously means ascribing a new meaning to the level of the individual actors. In spite of all the formal restrictions of the theory of games, this makes it possible to take into account that individuals do indeed act according to their social roles, and are therefore to a certain extent predetermined by the systemic rules - although these are stochastic - but at the same time they act according to their own interests too, which they partly define themselves, and thus they fill the scope that is always granted by social roles individually. Naturally, the theory of games is at least as formal as the foundations of a theory of social systems developed here; Therefore, "individually" must not be equated with "arbitrarily". In fact, the theory of games explicitly postulates the formal equality of

[84] Lansing and Kremer note somewhat ironically that this selforganized network is much more effective than the central plannings of the Indonesian government. After all, the reason why this net really works is of course that the peasants in the single villages have learned that they can individually prosper only if they cooperate with other units.

rational players. Nevertheless, the rational pursuit of their own interests, and the resulting behavior towards other "players," form a level that develops its own, partially independent logic, although the system as a whole plays a fundamental part in it, namely by setting the pre-forming frame for generally possible behavior. Strategic games "in general" can be investigated in the same way as the previously analyzed "pure" systems, regardless of comprehensive systemic contexts - with exceptions though, which will be dealt with below. In particular, it is of course possible to comprehend the development of strategies of games itself as a specific system dynamics. We have already done that in broad outline in section 3.2. with the iterated prisoner's dilemma.

This has to be explained in more detail. On the one hand, we do not intend to put forward any anthropological assertions like, e.g., people always behave "rationally" and "egotistically". Neither of these claims could be maintained empirically, and they are much too vague anyway. For example, it can be perfectly reasonable and thus "rational" for an individual not to behave egotistically in a particular situation, but instead to look for a strategy that is, e.g., orientated to the saving of another person (cf. e.g. Esser 1991). Our concern is that a system works as a "preforming frame" in three ways: first, the system determines *which* roles do or do not have to be realized primarily or exclusively by strategic action - e.g., the role of an employer vs. a parental role. Second, the system determines the "pay-off matrix", i.e., the amount of the bonus for successful or unsuccessful action (see below). And finally, the degrees of freedom of the individual strategy *within* the role is determined in the system, i.e., the radicalness that an individual actor can summon up to pursue his subjective interests. These are sociological questions, and not psychological-anthropological ones; in this sense, a sociological systems theory is always systematically superordinate to a (sociological) theory of games. However, if the theory of games is taken as a more precise version of a sociological theory of action, there is no such sub- or superordination: asking for the systemic frame conditions of strategic action is as necessary as asking for its systemic consequences and regarding systems as the result of individual actions.

Hence, the rational pursuance of one's own interests can not occur only in the context of general system rules, but in return can also lead to changes of the system rules themselves. In this respect, Max Weber's famous study of the "Protestant ethic" (1981), which is with justification called classical, can also be read in the following way: rational players - the Calvinist citizens and early capitalists - had no other choice in pursuing their interests, namely maximization of profits as an indicator of their religious "choseness", than to change the system rules of the pre-

bourgeois society in order to be able to realize their strategic interests. The fact that these strategic players were subsequently determined in turn by the changed system rules is actually nothing other than the dialectical logic between different levels of social reality that has already been described by Hegel and Marx. This problem will be analyzed below with the example of an extended version of the prisoner's dilemma.[85] We might also couch the case in terminology borrowed from Luhmann: strategic actors' games realize specific interaction systems within greater systemic contexts; these strategic interaction systems, as outlined above, develop their own dynamics, and by that have an effect on the whole system, or on the actually superordinate subsystems, and subsequently are limited and controlled in turn by the superordinate system.

The central question by which the theory of games was attached to systems theory, as it were, was the problem of the development of cooperative vs. aggressive behavior; in the following, we will concentrate on this paradigmatic problem (see Rapoport 1988 for the various questions with which the theory of games that is particularly orientated to decision theory is dealing). The reason for this interest, which is shared by social scientists and biologists, is of course the fact that aggressive behavior always seems more favorable from a rational-egocentric standpoint - provided one can get up to it within specific power constellations. Hobbes' famous question of how the quasi-natural aggression of man could be controlled in terms of civilization, and his no less famous answer to it, obviously presupposed that aggression is a reasonable behavior from the individual's point of view, but not necessarily from the point of view of the species or a social system. As a consequence, the Leviathan of the state had to be created to end the "*bellum omnium contra omnes*". This is not the place to make a political comment on Hobbes' theory; it should be obvious anyhow that his solution is deeply unsatisfactory on logical grounds.

Biological formulations of the evolution of cooperative behavior can be reviewed accordingly: on the one hand, "cooperation is essential for evolution ... (yet) cooperation may be difficult to achieve in a classic Darwinian framework. Cooperators have to succeed in the struggle for survival with defectors, who by definition have a certain fitness advantage." (Nowak/May 1993, 35). The various explanatory attempts often bear an ad hoc character.

Axelrod, and other game theorists in collaboration with him, tried to find a different, namely mathematical answer. Axelrod's famous study

[85] In order to avoid any terminological misunderstanding it has to be emphasized that those social levels are completely dissimilar to the different logical levels of hybrid systems. The logical levels are "layerings" of the rule systems, not of the social systems.

(1984) was to demonstrate that cooperation almost necessarily arises out of specific conditions, which can be mathematically formulated, in fact with justification, especially when all actors involved behave in a rational-egotistical way. The strategy of TIT FOR TAT (TFT), which was already described in section 3.2., is a cooperative strategy, in that a TFT player always tries to cope with the other players and opponents, respectively, on the basis of cooperation. Only the encounter of someone else's aggression can get a TFT player to *respond* with aggressive strategies himself. The noted successes of TFT in Axelrod's two computer tournaments and his ensuing mathematical analysis show in fact that a TFT strategy is obviously superior to purely aggressive strategies - with particular restrictions of course. Before analyzing this in more detail we are going to take a closer look at a specific version of the iterated prisoner's dilemma by Nowak and May (1993); it is particularly suitable to show the connection between the theory of games and systems theory.[86]

Nowak and May orient themselves toward a modeling of the iterated prisoner's dilemma with a CA, which was already introduced by Axelrod (1984). The basic idea is that the spatial situation of strategic players can play an important part in the question of how the system as a whole, consisting of various players with different strategies, develops, and, in particular, on which conditions a dominance of cooperative or aggressive strategies, or to an equilibrium between them develops. We call a strategy "cooperative" if a player responds to cooperative moves of the other players immediately or after a short time with his or her own cooperative behavior; strategies are "aggressive", by contrast, when they try to "exploit" cooperative moves of the opponents with aggressive moves.[87]

The CA of Nowak and May is two-dimensional, and its basic version works with a Moore neighborhood. The cells can take one of two states,

[86] Obviously, the Balinese peasants in the above example also had the alternative between a behavior with no consideration for their neighbors, that is, aggressive, and cooperative behavior. It was mentioned that Lansing and Kremer did not analyze the network example in a games theoretical but in a systems theoretical way. Accordingly, one might reconstruct the example as well in the form of the iterated prisoner's dilemma; apparently, the decision on whether to choose the theory of games or systems theory is just a question of perspective here - "from below" or "from above".

[87] In section 3.2. we showed graphs of proceeding and trajectories of a system of strategic players that orientated themselves to the different strategies' degree of cooperativity. This was defined as the number of possible cooperative moves within a total strategy. Thus, a strategy is "cooperative" if its degree of cooperativity exceeds 50 %, otherwise it is aggressive. As regards content, this amounts to the above informal definition.

namely 0 for the strategy D and 1 for C. The pay-off matrix is changed in comparison with Axelrod's standard version, and has the following form:

	C	D
C	1	0
D	b	0

where b is a variable whose influence on the system dynamics has been analyzed by Nowak and May. The basic idea behind that is immediately plausible by intuition: the decisive factor for the dominance of cooperativity or aggression within a system is how "rewarding" aggressive strategies are. (Of course, it would also be possible to keep b constant and allow the pay-off of two cooperative players to vary instead.) Now every cell "plays" with each of its eight neighboring cells and with itself. The pay-off of for a cell is the sum of all results of all moves with the eight neighboring cells and with itself here. For example, a cell in the state C, whose neighborhood exclusively consists of C cells, gets a pay-off of 9, a D cell, in contrast, gets a pay-off of 8 * b in the same neighborhood.

In the next "generation" of the game, that is, at the next run of the CA, a cell turns into the state that corresponds to the state of the neighboring cell that received the highest pay-off in the previous game (the cell itself is of course included in this computation). Apparently, this not only depends on the cell's game against its eight neighboring cells and against itself, but also on the games which the eight neighboring cells played against their actual neighboring cells. Therefore, an "extended Moore neighborhood" of a total of 24 cells, plus the cell itself, has to be taken into account, which results in 2^{25} different transitions. If we assume that the extended Moore neighborhood in the first example of a C cell consisted of C cells as well, then all of the 9 possible cells have attained the pay-off = 9, and our C cell remains in its state. If in the case of the D cell the extended Moore neighborhood also consists of C cells, then the D cell attains a pay-off of 8 * b, as was mentioned, and the C cells of the simple Moore neighborhood each attain a pay-off of 8. Here it seems to depend only on the size of b whether the D cell changes its state or not. If $b < 1$ then the D cell changes into the state C; if $b > 1$ it remains in the state D.

Nowak and May chose a CA modeling mainly because they wanted to study the influence of spatial relations on the development of specific strategies; they started out from the certainly realistic assumption that neither all players play against all others in greater systems nor are the actual partners selected at random. One of these two presuppositions underlies most analyses of the iterated prisoner's dilemma with greater

populations. Being theoretical biologists, Nowak and May assumed, by contrast, that only players within a spatial neighborhood (can) possibly play against each other, and that this can be optimally modeled with a CA. In the next section we will be concerned with an application of this thought to social spaces; in fact, the Balinese irrigation network only evolved from the local cooperation of neighboring social units as well.

However simple the above listed rules of the game may be, the real system dynamics is unpredictable, as Nowak and May emphasize. Of course, this comes as no surprise for those who know about the complex dynamics that CAs can develop; let us therefore look in more detail at some results that were attained in the experiments with a different b.

It is intuitively plausible that the chance for aggressive strategies to win through within the system is greater the higher the value of b is. In particular, it is trivial that it is necessary that $b > 1$, otherwise every D cell, provided there is just one C cell in its neighborhood, would turn into the C state (in the case of a cell's game against itself not being computed, there would have to be two C cells). On the other hand, it is not intuitively comprehensible how great b has to be in order to (a) get more D cells than C cells or even only D cells; or (b) get approximately the same number of C and D cells; or (c) get mainly C cells (if there is a D cell at all and if $b > 1$, then that cell will only vanish on very specific conditions). In other words: b plays the part of a control parameter whose size controls the dynamics of the system.

Nowak and May analyzed the variations of b within the area $1 < b < 3$ (with $b > 3$ the dynamics of the system no longer changed). They summarized their main results with regard to b in the following way:

> "The above analysis suggests the existence of three classes of parameter regions:
> (i) If $b < 1.8$ then only C clusters can keep growing.
> (ii) If $b > 2$ then only D clusters can keep growing.
> (iii) If $1.8 < b < 2$ then both C and D clusters can keep growing." (loc. cit., 39)

If we translate these results into the terminology of complexity classes (which has since become well-known), which is what Nowak and May do as well, then b shows itself as a control parameter that generates our well-known different system dynamics at different values: if b is close to 1, the system behaves as if it belonged to Wolfram class I; values of b between 1.2 and 1.8 result in longer periods, and thus in a behavior of class II. The periods get longer the more b approximates the value 1.8, and between 1.8 and 2 the system tends toward chaotic behavior. If b keeps on growing the complexity goes down again, which distinguishes b from the "monotone" parameters like P, Z, v, and λ of course; at $b = 3$ it

attains a behavior of class I again, though with a clear dominance of D, that is, aggressive strategies.[88]

However, in terms of a "parameter theory" we have to point to the fact that the given values of b apply only within the above pay-off matrix. If we set the other three values of the pay-off matrix differently, with unchanged relations, then the relevant values of b would change; the basic structure remains the same, however, in the sense that the three mentioned "parameter regions" always arise. In contrast to our well-known control parameters, whose magnitudes are "universally" valid with the mentioned restrictions as it were, the values of b exclusively refer to a specific pay-off matrix.[89] If one starts out from a fixed pay-off matrix, then the values of b can be directly applied to corresponding values of P or λ, so that it is possible to give a sufficiently adequate prediction of the resulting system dynamics.

Before we appreciate the results of Nowak and May, as well as similar results of Fogel (1993), who also analyzed systems with different pay-off matrices comparable to CAs, with more regard to their content, we would like to summarize some of the most important game theoretical results of the two "classics", Axelrod and Maynard Smith, in order to elucidate the matter:

An analysis of the possibilities of the "evolution of cooperation" that is based on the assumption that strategic actors are only interested in their own advantage, and on the assumption of prospects of success that correspond to the pay-off matrix of the iterated prisoner's dilemma, really leads to the fact that the Hobbesian answer is not cogent for mathematical reasons alone. It can in fact be shown that a relatively small number of cooperative actors, according to a strategy similar to TFT, is already sufficient for these strategies to quickly win through and assert themselves against purely aggressive strategies. In particular, it turns out that on the one hand purely aggressive strategies are not "evolutionary stable" (Maynard Smith 1982), i.e., they cannot assert themselves against other strategies; and on the other hand that "mixed" strategies like TFT usually have an advantage in the medium- and long-term - "purely" cooperative strategies like "always turning the other cheek" are not evolutionarily successful either. Thus, the successes of TFT in the mentioned computer tournaments are no interesting accident that is only be explainable from the specific conditions of the

[88] Obviously, b is a "general" control parameter (see 3.1.2), which determines (like P or λ) the relative frequency of realizing certain cell states.

[89] Without having analyzed it in more detail yet, we (Jörn Schmidt and J.K.) assume that the absolute values of b directly depend on the determinant of the pay-off matrix and not just on the matrix' single values. Nowak and May do not indicate such relations but only report experimental results with the mentioned matrix.

tournaments, but instead point to a deeper logic of strategic situations: the Hobbesian problem can be resolved by the fact that its premise is not valid. A *bellum omnium contra omnes* does not have to be ended by an almighty Leviathan, but instead ends itself, so to speak, by the fact that at the occurrence of "mutations", i.e., mixed cooperative strategies, the purely aggressive actors prove to be evolutionarily unstable, and "die out" in favor of the more successful new strategies, or become mixed cooperative actors themselves. The experiments with GAs in section 3.2. show that this is also true on essentially changed conditions: in these highly stochastic systems, mixed strategies always win through in the medium-term; if purely or predominantly aggressive strategies reappear by mutation, they quickly vanish again because of the mixed cooperative strategies' superiority.

The evolutionary stability of mixed cooperative strategies becomes particularly clear from the difference between the models of Axelrod and our models. As described in section 3.2., we took a vector of 63 binary components for the coding of a strategy; the figure 63 results from the fact that the final 5 moves of the opponent are (or at least can be) taken into account each time. In comparison, Axelrod's model (1987) consists of binary coded vectors of a length of 70; here a player can look back on the last three moves. However, he can do so not only on those of his opponent but also on his own moves, so that the opponent's behavior can also be evaluated as a response to his own behavior in principle. This results in $4^3 = 64$ possibilities; Axelrod adds another 6 components for three "hypothetical" initial moves, so that the vector consists of a total of 70 components.

Although the two models are technically very similar - in fact we got our inspiration from Axelrod's model - they obviously differ in one assumption on each's part, which might have quite serious effects in reality: in Axelrod's model the opponent's and the own actions are considered as interdependent in principle, and both have to be taken into account when the next strategy is chosen. This is different in our model; here only the opponent's actions are called into play for the selection of one's own strategy: if they are mainly aggressive, the response is according, that is, it is also aggressive, if possible; if there are other strategies, the behavior is again different, etc.

If this difference is translated into common psychological and/or political terms, then an opponent is only judged by *his own behavior* in our model; he "is" aggressive, so to speak, if his last five moves were aggressive, and so forth. In Axelrod's model, by contrast, the opponent's behavior is judged as *a product of his and one's own behavior*. An opponent "is" not aggressive, but he acts aggressively or cooperatively in

relation to his own strategies. Of course, an opponent can "be" aggressive here, too, in the sense that he permanently chooses aggressive moves with no consideration for the moves of the other, and regardless of his own behavior. But since this is not an evolutionary stable strategy - see above - such an opponent will harm himself in the long-term (and others in the short-term as well of course).

Evidently, the basic assumption in Axelrod's model corresponds rather well to what might be regarded as a "reasonable" policy. Both the course of the Cold War and the history of political "arch-enemies" has repeatedly shown how incomplete and unfruitful it is to judge the actual opponent *exclusively* by his own actions. A nation "is" not aggressive, but is more or less so in the context of its opponent's actions, not least because of its experience with the previous actions of the other players. If this basically rather simple insight would have been considered more often, innumerable debates would have been unnecessary - from those concerning the fundamentally "belligerent" nature of the Germans or the French, to Reagan's Evil Empire. Since such an insight presumably requires a different kind of politics, one might expect that this should be confirmed by different results in experiments with the two models; in fact we had deliberately constructed a divergent model in order to analyze the differences we were expecting.

However, it is characteristic that there were no significant differences *in the general trend.* In Axelrod's model, as well as in ours, mixed strategies regularly predominated in the medium-term, unless it came to playing against a purely aggressive strategy - "defect all". The latter simply results from the fact that a purely aggressive strategy is just the only reasonable thing to do when playing against a single, purely aggressive opponent; the genetic algorithm quickly figured that out of course. In all other cases, and particularly in playing each against each other - see above, section 3.2. - the mixed strategies postulated by Axelrod and Maynard Smith came out as the evolutionarily successful ones in both models. Hence, we can interpret these results in the way that in terms of the theory of games it makes no difference if an opponent is judged by his behavior alone, or also by the interdependence between one's own and the opponent's behavior: the most favorable strategy *for oneself* is all but the same in both cases. In other words: a strategic interaction system will always show almost the same dynamics, predominated by mixed strategies, regardless of how simple or differentiated the actual opponents' evaluations of each other are. The very fact that the successful strategy of tit for tat considers only the last

move of the opponent without any reference to its own moves clearly illustrates that.[90]

Of course, these results do not dispute the general reasonableness of a differentiated assessment of opponents in the above sense. We should keep in mind that all the reported results are valid only under the premise of the situations of decision being *iterated*, so that the cooperative mixed strategies are able to develop and win through against the aggressive strategies. All the results tell us that cooperativity is only rewarding in the medium- and long-term and, as Keynes said, "in the long run we are all dead". Many situations of decision on the political stage offer only one chance to decide, and in that case a too undifferentiated assessment of the opponent can literally have fatal consequences at the first instance. Thus, game and systems theoretical analyses are often not suitable for single situations, but their value lies in the general comprehension of the dynamics that result from strategic behavior. The analysis of pure game systems can provide important information for the classical sociological and evolutionary biological question of how the evolution of cooperativity can be explained; the analysis of single situations of decision is really something different, even though the general usefulness of TFT strategies, which was frequently emphasized by Axelrod, surely has its justification in very concrete situations.

The iterated prisoner's dilemma has been studied again and again in numerous varieties, and we can only briefly refer to them here. In particular, we should mention the experiments of Lindgren (1992), who also operated with GAs; here the players received partly incomplete information on the opponent's strategies ("noise") which resulted in highly complex dynamics. Lindgren, who studied the iterated prisoner's dilemma from an evolutionary biological perspective similar to that of Maynard Smith, showed that incompleteness in particular is an important factor in the development of complex evolutionary paths. Yet more variants were realized by Stanley et al. (1994): in their version the players have the possibility of rejecting or accepting other players as potential partners. The basically quite simple dynamics in the basic models of Axelrod, Maynard Smith, and ourselves also become more complicated indeed through the introduction of such additional conditions, which are more adequate to reality; however, even in the more complicated models the evolution of cooperative behavior still

[90] In 1987 Axelrod reports his results a little bit unclear and too simple so that we first thought to have obtained significant differences both between our model and that of Axelrod and in repeating Axelrod's experiments with his own model. However, after Axelrod kindly put more exact graphics of his results at our disposal it turned out that these were primarily semantic differences. Once again we would like to express our thanks to professor Axelrod for his cooperativity.

arises as a characteristic feature of the system, while the single actors behave egocentrically.

Unfortunately, these solutions of the Hobbesian problem, however acceptable they may look, are incomplete. Here as in many other cases of deliberately simplified modelings - see the remarks on control parameters above - the famous skepticism of Sporting Life from "Porgy and Bess" applies: "It ain't necessarily so." We can easily see that by taking a closer look at the results of Nowak and May, and those of Fogel:

It was emphasized again and again that the logic of the - simple or iterated - prisoner's dilemma ensues from the pay-off matrix. If the latter is kept constant, as it usually is, according to the basic version of Axelrod and Rapoport or the starting matrix of Nowak and May, then we get the reported results with different models and techniques of simulation. If, e.g., the b value is systematically changed in the matrix of Nowak and May, we get completely different system dynamics, as was shown above, and cooperativity especially has no, or only a small, chance. Therefore, the results just mentioned, which in fact give rise to optimism, are only valid in very specific classes of "relations of reward and punishment," and in other classes they are not. As a consequence, Hobbes' initial question cannot be answered by simply referring to the medium- and long-term superiority of mixed cooperative strategies to purely aggressive ones; besides that, one has to explain on which pay-off conditions cooperativity evolves with a certain necessity; on which conditions both aggression and cooperativity, will evolve; and at what point cooperativity has no chance even in the long run (cf. Fararo 1989). Thus, sociologically the question of the evolution of cooperativity is connected with the empirical question of what the pay-off matrices look like in human and animal societies. The fact that, e.g., chimpanzees can show a high degree of aggressive behavior towards other chimpanzees (Goodall 1986) indicates that we are obviously dealing with empirical constraints here as well, as in the case of human beings.

This can even be intensified according to the famous "principle of St Matthew" - whosoever hath, to him shall be given. In the experiments of Nowak/May and Fogel, the single pay-off matrices had been kept constant during the experiment. This is not quite realistic though, since social systems, as has been stressed again and again in this study, can in principle vary all of their rules; and a pay-off matrix is actually nothing other than such a rule, or a parameterized property of the respective rule system. In this context, an orientation to the principle of St Matthew would include the possibility - and probability - of the pay-off matrix being changed in favor of the actually predominating actors, while the system dynamics are in progress: if cooperative or mixed cooperative strategies predominate in the system, the b value is reduced in the matrix

of Nowak and May; if aggressive strategies predominate, the b value is increased; and if cooperative and aggressive strategies are equally matched, the b value remains constant. Just a single glance at history shows that this permanently happens in social systems: if new aggressive players appearing after periods of peace are rewarded so that their aggression seems to pay off, then these aggressive actors will do anything to keep and even increase their advantages. The continual escalation of violence that could be observed lately in the former Yugoslavia can be explained at once according to that principle; the aggressive actors (whoever they were at the beginning) continually increased the rewards for their aggression themselves.

In order to examine these considerations in more detail, we extended our own model in the following way: the system " each against each other" (see above, section 3.2.) starts, as usual, with the pay-off matrix introduced by Rapoport and Axelrod

	C	*D*
C	3	0
D	5	1

and thus b = 5 in the terminology of Nowak and May, and with strategy vectors generated at random (it is also possible to instruct the program to generate an equal number of mixed cooperative and mixed aggressive vectors, but that does not make much difference in the medium term). After one "tournament", i.e., after about 100 moves that any strategy performs against any other, 20 new vectors are generated by the GA in the way described above. Before the next tournament starts, the program checks whether there are more cooperative or aggressive strategies - "players" - in the system. If there are more aggressive strategies, b is increased by a specific value, e.g. 0.05, and the reward is reduced for two cooperative players; in the opposite case b is reduced and the cooperative players' reward is increased; if the number of cooperative and aggressive players is equal, the values stay as they were at the start. Then the next tournament is played with the pay-off matrix of the end of the first tournament, the pay-off matrix is set again, and so on until the whole experiment is finished (usually after about 120 "tournaments").

It seems plausible that with such a variable pay-off matrix, a starting advantage can hardly be made up. For example, if there are by chance more cooperative strategies at the beginning which pay off successfully in the first tournament, then this advantage is consolidated further by the variation of the matrix. In such a case it can be expected that the whole system permanently moves towards fully cooperative strategies, and that

aggressive strategies, which come into the system by mutation, are eradicated as being evolutionarily unstable. Starting advantages of aggressive strategies would have a corresponding effect.

Interestingly enough, things are once again not so simple. Our experiments showed that the aggressive strategies were favored to a highly significant degree, no matter whether the initial vectors were generated at random, or whether the system consisted of an equal number of cooperative and aggressive strategies at the start. The b value of the pay-off matrix was increased after the first tournaments with a virtually irritating regularity, which resulted in the aggressive strategies becoming more and more evolutionarily successful, as was to be expected. Even if the cooperative strategies had a starting advantage and b was reduced a little bit, this was usually insufficient; after a relatively short time - 15 - 20 tournaments at the most - this advantage leveled out again, and b began rising up to its maximum value of 7.

Looking back on the results of Nowak/May and Fogel, we can quite easily explain this outcome, which might seem somewhat surprising at first: apparently, the gain in the case of aggressive behavior against a cooperative player - the b value - is so high compared with the reward for mutual cooperativity that *for the short-term* the aggressive players predominate. Once again, one can discern the importance of the *iteration* of the strategic game here. However, as soon as aggressive players predominate even for a short time, they are immediately able to change the pay-off matrix in their favor and by that stabilize their advantage for the long-term too. This is not true in the opposite case though, because the advantage of mutual cooperativity only has a medium- or long-term effect. Thus, a pay-off matrix with b values that strongly favor aggressive behavior in the short-term necessarily leads to an establishing of aggression as the predominant basic pattern in stochastic systems too.

Of course, from these results arises the question of how high the *initial* b value has to be in order to obtain the described outcome, and at what point other effects appear when the b value is reduced, respectively. So we systematically reduced b, that is, we let b attain different values in the interval of 3 (the reward for mutual cooperativity) and 5. If the other values of the above matrix are kept constant, then the following result ensues, whose absolute values apply of course only with regard to this matrix: if the initial value of b is at about 3.2 the cooperative players have a realistic chance to make use of starting advantages in the long-term. At higher values of b, in contrast, the above analysis is true: the short-term advantages of the aggressive strategies are so predominant that even starting advantages of the cooperative players cannot influence the medium- and long-term course. The advantage that can be granted to aggressive behavior in comparison with mutually cooperative behavior

may only be small if cooperative behavior is supposed to have a chance at all.

Coming back to Hobbes once again, after all we said before, the answer seems to be somewhat more pessimistic than the one that Axelrod in particular gave about the evolution of cooperativity: if in a "strategic interaction system" the advantage of aggressive behavior is relatively great in relation to cooperative behavior at the beginning of the history of the system - technically speaking: $b > 3.2$ - and if the system is variable in the sense that the actors can use a short-term predominance to increase their advantages, then cooperative behavior has not got a chance even in the medium-and long-term. This is also true at high mutation rates, which regularly bring cooperative actors into the system; however, according to our results these can neither hold up at the beginning, when the pay-off matrix is still favorable for them, nor afterwards. In such a system it makes no difference whether the players "actually" use a relatively cooperative TFT strategy. The predominance of aggressive strategies works in such a way that only the aggressive moves of a TFT strategy, which is in fact a mixed strategy, can actually have an effect; this becomes very clear in our experiments: at the end of a tournament with very high b values for the pay-off matrix, the system does not entirely consist of purely aggressive strategies, but instead there is a great number of mixed strategies as well. However, the results of the tournament show that only the aggressive moves are really carried out.

If we want to avoid the Leviathan as the only solution of the cooperativity problem, it seems that only the following possibilities remain: (a) We assume that early human communities either have a constant pay-off matrix, or that the b value is so low that the inevitable developments that were described above can be avoided. This is an empirical question, and there is probably no definite answer to it. In fact, aggressive behavior did pay off quite often in history, which makes it rather improbable to escape the general assumption of a variable pay-off matrix. Hypostatizing a social primordial state, by contrast, in which the pay-off matrix was both constant and favorable for cooperativity, makes one think of Marxian early communism, and other postulates of golden ages; as far as we know, the present results of ethnology and biological behavioral research do not back up this assumption. Anyway, such a hypostatizing can never be proven.[91]

[91] One may of course assume that cooperativity between *relatives* was predetermined by genetically fixed patterns of behavior, which the "kin selection theory" is known to postulate. However, there still remains the corresponding problem of how it was possible for cooperative behavior between non-relatives and foreign societies to develop and pay off under the above circumstances.

(b) The only way out of this dilemma probably lies in pointing to a categorical deficiency of this and other game theoretical procedures: all of the above mentioned game theoretical models (and practically all others too) contain the tacit presupposition of the feasibility of a purely *intrasystemic* determination of success and failure, respectively, of single strategies. The pay-off matrix only determines rewards and punishments for the single players in relation to each other, regardless of what this means for the system as a whole. The only reason for the fact that mixed cooperative strategies often win through is that this is to the advantage of the single actors in the medium- or long-term. If, on the other hand, aggressive behavior is more useful to them, and if they can increase this usefulness, then they will do just this. The fact that a realization of a long-term predominance of aggressive strategies is less favorable for the system as a whole than a predominance of cooperative strategies is left out of consideration. This can be seen, e.g., from the history of the Thirty Years' War: the short-term advantage of individual aggressive behavior seems to have been so great for most of the actors that they tried to realize this advantage; the fact that the situation of the social system "German Empire" thereby became worse and worse as an "unintended consequence" was only of interest in the end, when most of the actors had practically ruined themselves (on this example see Knorr-Cetina 1981).

So, strictly speaking, a strategy has to be evaluated several times: first, the advantages and disadvantages for the single actors in relation to other actors must be considered of course, like in the pay-off matrix. Formally speaking, however, these are ordinal values that provide no direct information on whether that behavior pays off "absolutely", that is, e.g., by a real increase of one's own wealth. A second evaluation would have to grasp just this, namely by judging a strategy not only by its relation to the successes of the other actors, but also with regard to what it brings in for the single actor in relation to *what can be achieved in principle*. Along with this, a strategy is then evaluated with regard to what it achieves for the system as a whole: when all strategies are taken into account, the difference between a strategy and what can be achieved in principle turns out to be the difference between the system itself and the environmental requirements that require a particular performance of the system as a whole. This is informally the case with the frequently mentioned Balinese peasants: a short-term egotistical exploitation of the water resources with no consideration for the neighbors may indeed increase the village's wealth. However, the equally egotistical (aggressive) reaction of the neighbors would eventually lead to a breakdown of the entire irrigation system, and thus the single actors' welfare follows from the welfare of the system as a whole; hence, the

"systemic value" of a strategy with regard to a neighboring village depends on how well the system as a whole can fulfil the conditions of its ecological environment.

This can be stated more precisely in an extended model: a system value calculated as the (arithmetic) mean from the values that all 20 strategies (= players) have obtained in one tournament is assigned to the above system with variable pay-off matrix. The assumption here is of course that social interactions are not only significant for the respective actors, which is tacitly presupposed for the most part in the theory of games, but that they also mean something for the system. The fact that systems define themselves according to the interactions of their elements, and that this also is their only possibility for realizing their adequacy with respect to specific environmental requirements, should make this assumption immediately plausible. Since the system is stochastic, a mean value orientated definition of the system value or a definition according to the total sum suggests itself; taking only the best 5 or 10 players would necessarily bear the permanent danger of statistical "runaways".

The environmental requirements are now represented as a scalar and not as a vector, because the system value is a scalar as well. The level of the environmental requirements is determined experimentally, namely by the fact that the system is still able to fulfil the requirements with a slight predominance of aggressive strategies. Now if aggressive players are rewarded at the beginning and succeed in changing the pay-off matrix in their favor, as it is usually to be expected, the system will diminish its total value. If the value falls too much below the environmental requirements, the system gradually increases its *mutation rate*: this follows the assumption that if the situation of the whole system, and along with it the situation of the actors of course, constantly deteriorates, there will be a growing readiness of the actors to give cooperativity a chance, so to speak. This happens in the sense that the crossover is carried out first according to the well-known selection principle, and subsequently the new vectors are subjected to mutation (in fact, the selection first of all favors the aggressive strategies). What follows is the analysis of the frequency of aggressive and cooperative strategies respectively; the determination of the pay-off matrix; the new tournament; and the computation of the system value. If the latter is still too low, the mutation rate is increased further (when aggressive strategies are predominant there is of course a high probability that the increased mutation leads to more cooperative strategies) and so forth, until a "satisfactory" system value is achieved, i.e., until the strategies average high values. As soon as that is the case, the mutation rate is reset to its initial value again and the system proceeds in its development. The

interesting question here is of course how long and how radical the system has to change its mutation rate in order to get out of the self-produced misery, and which b values are attained by this (in fact, the b values do not directly arise as independent variables here, but as dependent on the system values). It could indeed be shown that cooperative behavior has a real chance in these conditions, even on the assumption of a fully variable pay-off matrix. Ralph Kier will present detailed results on this in his doctoral thesis.

We can see from the logic of this model that it is indeed possible to escape the necessity of a Hobbesian Leviathan even under the above difficult conditions, i.e., the furthering of aggression according to the principle of St Matthew. Of course, this requires the presupposition of the system being sufficiently variable, which not only implies its capability of changing the pay-off matrix, but also of changing the single actors, i.e., their interaction strategies, by means of a "willingness to learn," until the system achieves favorable values "by itself" again. The above examples - from the Balinese village to the Thirty Years' War - might equally well be interpreted in the following way: the actors changed their behavior because of bad system values until these values had again become fairly adequate.

Of course, this does not necessarily work. A system with a minor willingness to learn, that is, a not particularly variable system, might just as likely ruin itself completely before the variability is sufficiently increased. In the case of the civil war in the former Yugoslavia, the pay-off matrix apparently had to be changed from outside, namely by NATO, because the system was not able to do it itself. This corresponds with our own results: a mere small increase of the mutation rate had practically no effect on the system as a whole; selection and crossover favored the more aggressive actors again and again, so that the system could not get out of its misery.

Before concluding our reflections on this subject, we would like to describe one last model in which we (Jörn Schmidt and Jürgen Klüver) attempted to combine game and systems theoretical aspects of social systems. It is the strategic interaction of hybrid cellular automata; the program has the beautiful name SOCVAR = variable societies (SOCietates VARiables).

The ordinary strategies for the iterated prisoner's dilemma (and not only these) are so-called "look back" strategies. This means that each move orientates to the opponent's *previous* behavior, as is especially clear in the example of tit for tat. This orientation can be as simple as with tit for tat, or much more subtle, as in the model of Axelrod (1987); the fundamental logic is always the same. There are in fact strategies that try to determine the opponent's move that will presumably follow on the

basis of probability assessments, and then derive their next move therefrom, such as DOWNING, which has been described by Axelrod (1984, 34); however, these are exceptions. Anyway, in those cases one might speak of a "look forward" strategy.

In real decision situations it is certainly often the case that the strategic actors do not (or can not) content themselves with a look back strategy, but must also take their opponent's presumable moves, and *in particular the probable effects of his or her own decisions*, into their calculations. This is certainly true for classical situations of decision, such as chess or war. So, both the probable moves of the opponent and, above all, the consequences of one's own actions are assessed. In other words: strategic actors *anticipate* the possible or probable future that results from their actions and those of their opponents, and act on the basis of such anticipations. The fact that successful strategies in the iterated prisoner's dilemma do not need to do this does not lower the value of this game theoretical paradigm, though it points to the fact that game and systems theoretical analyses of situations of decision have to go beyond this pure, basic model. In the theory of games this is of course taken into account (cf. Rapoport 1988).

In a logical sense hybrid systems like, e.g., SOZION seem to be able to anticipate. The fact that SOZION "increases itself twenty-fold" enables it to test in anticipation which rule system is the best, i.e., which decisions for or against particular rule systems it shall make. It suggests itself to use this ability for strategic interactions in order to check the prospects of success and the evolution of such systems.

The program SOCVAR consists of four hybrid CAs of the type SOZION with the above described rules of differentiation. Thus, each CA is "controlled" by a GA that generates a total of twenty possible alternatives for each CA and performs the selections on the basis of the realized system values. The program starts with four CAs, each of which has 100 "members", namely cells in the state rural population. These clusters are spread homogeneously across the screen so that each has the same amount of free space for its own growth.

The spatial homogeneity is important because previous experiments with an earlier version of SOCVAR have shown again and again that clusters with more space than others were regularly superior. This was also true if those CAs did not develop as well as the smaller ones. Sheer size is an extremely advantageous factor in itself, which can be gathered, e.g., from the history of Russia or China. If systems can spread in a free space so that they are much larger in comparison to the competing systems they encounter, then the larger systems always have greater chances *provided that their variability is approximately the same.*

Consequently, we had to keep the spatial conditions equal, since one important point of our analysis of SOCVAR was to test the influence of variability on the systems' strategic success.[92] Thus, all four systems have an equal chance of spreading in the free space at the start of the program. This is realized by a growth rule analogous to SOZGEN. The inner processes of differentiation in all programs pass off in the same way as in SOZION and SOZGEN; this is stimulated by an environmental vector which guarantees a minimum of socio-cultural evolution at the beginning of the program.

If the single systems encounter each other now after a period of growth, rules of interaction set in that make the whole thing seem like an extended form of the iterated prisoner's dilemma: a system has three possibilities with regard to another system, namely: (a) "Coexistence" - the other system is ignored and nothing happens except for the fact that the system's own growth stops at that place; (b) "Cooperation" - something like an economical and cultural exchange is offered, which the other can accept or decline; (c) "War" - the system's own cells that are in contact with the other system are transformed into "warrior cells," which will destroy a cell of the other system in the next run. This creates a free space to which the warrior cells move in, whereupon the belligerent system can grow into the place of the warrior cells. As a result, the belligerent system can grow at the other system's expense.

In order to let the subprograms decide for the systems what is best for them, they are given a pay-off matrix, which now has to consider three possible strategies of course. Let D be the symbol for the option of war, C the option of cooperation, and E the option of coexistence, then we get a 3 * 3 matrix for two systems:

	C	*D*	*E*
C	*a*	*b*	*c*
D	*d*	*e*	*d*
E	*c*	*b*	*c*

We must consider the following here: a = 2 * n, for n being the set of *mutually* cooperating cells. In fact, it is of no use to a system willing to cooperate if the cells of the other system do not cooperate as well; this can only be done by cells of particular types (peasants, merchants, priests), however. Cooperation with a as the reward only evolves if two cooperating cells of the two systems meet. For m being the set of the

[92] In future experiments we are going to test to what extent spatial advantages can be compensated by a higher variability. The history of European culture obviously shows that geographically and demographically smaller units are perfectly able to make up for these disadvantages by means of their high variability.

system's own cells that are destroyed by warriors of the other system, b = - m. (Strictly speaking, b results in dependence on how valuable the destroyed cells were to the system; we can abstract from that here however.) Since cooperativity cannot occur with a system that ignores the other system, c = 0.

The determination of d is a bit more complicated: systems do not produce warrior cells before they encounter another system; the system value exclusively depends on the number of "productive" cells for the system. Warriors are in fact economically and culturally "parasitical", and so *in the short-term* the production of warrior cells is counterproductive for a system, because they are subtracted from the system value. Only if the own warrior cells have created new space for the belligerent system to grow into, the newly created "productive" cells are added to the system value according to the standard formula, i.e., the system value increases by the number of newly generated productive cells; before that it decreases for the time being. So, in contrast to the standard matrix of the iterated prisoner's dilemma, aggression does not provide an immediate advantage here.

The value of a is not determined from the start either. If two belligerent systems meet, there are two possibilities, which do not contradict each other: (a) Two warrior cells meet. One destroys the other by chance, so that after a longer clash none of the systems has an advantage, but only the disadvantage that results from the creation of warrior cells. (b) The warrior cells meet "peaceful" cells of the other system. In this case, e results for each of the two systems from the number of spaces that are occupied with own its cells, minus the cells that are destroyed by the other system. Thus, neither d nor e are constant values, but instead they change from move to move so to speak; in such a situation the system with the higher generation rate of warriors certainly has an advantage. It is easy to see, though, that a system that runs into an aggressive opponent can only react with aggression itself after all, like in the case of the iterated prisoner's dilemma. *Whether* it does so, however, is hard to predict.

The rest of the matrix should be apparent; naturally, a coexisting system is exploited just as a cooperative one by an aggressive system, and it is also obvious that according to the definition, the pay-off of coexisting against cooperative or other coexisting systems is always zero.

Regardless of the fact that the above matrix is a bit more complicated than that of the standard version of the iterated prisoner's dilemma - primarily by introducing the time factor - its basic structure is indeed similar to that of the prisoner's dilemma: aggressors have advantages over peaceful systems; cooperativity is highly valued but rather unlikely

because it always requires a second cooperative system, as well as the local presence of cells that are able to cooperate; that is why on the whole it is definitely at a disadvantage with aggression. A new element is the inclusion of coexistence; it corresponds to the possibility of refusing partners introduced by Stanley et al. (1994). However, in Stanley et al. this is not part of the pay-off matrix, but instead a player always has to look for a new partner after a refusal. Our matrix is more realistic in this point, since ignoring coexistence is undoubtedly a frequently chosen strategy - just think of the shield policy of the feudal Chinese and Japanese cultures.[93]

With regard to their optimization strategies, the systems orient to the system values of the actually best system. Thus, there is a permanent compulsion to compete, even for the coexisting systems, since all systems are always in deficit except for the best one. In contrast to simple strategic players, the systems have a twofold problem here, as can be easily seen: on the one hand, they have to develop a strategy with regard to the three other systems; on the other hand, they must resolve their "domestic affairs", namely consider their own social differentiation and adapt it to the prevailing requirements. Since the values of all systems are constantly changing, the system finds itself in the situation of a "moving target": every good solution it has found may have become obsolete again by virtue of the behavior of the other systems by the very next time step.

It is obvious that the variability of the system is extremely important in such a situation. Since all the CAs were of the same type as SOZION, they all had the same μ value, so that we could focus our experiments primarily on the success of systems with different values of κ, that is, for the sake of simplicity, on different values of h and r. All systems use an anticipating look forward strategy. The results reported here refer to systems with "anticipation size" 1: the systems test possible alternatives over one run of the CA before adopting the CA that achieved the best system values. The other systems are each analyzed with regard to their system value and their actual behavior, that is to say, their last move; the anticipation works on a "*ceteris paribus*" principle, which means that a system assumes that the other systems will repeat their previous interaction behavior at the next time step. On the basis of this assumption, it turned out that an increase of the anticipation size did not

[93] We (Jörn Schmidt and Jürgen Klüver) had developed this program before we studied the theory of games and the iterated prisoner's dilemma systematically. Afterwards we were surprised that we had apparently developed an extended variant of the dilemma. In our opinion, this shows that the situation of the dilemma, which might appear quite artificial at first, is in fact quite realistic (for "real" examples see especially Axelrod 1984 and Maynard Smith 1982).

bring any advantage worth mentioning for singular systems: the other systems usually changed their behavior so promptly and frequently that anticipations over more than one step quickly proved to be relatively worthless.[94]

As we said before, the κ value of a system, which is important here, depends on the number of uses of the GA - the h value - and the number of components allowed to be varied in the GA vector - the r value. The following values, e.g., produce a very illustrative result:

In system 1, h = 0.5 and r = 0.25; this means that the GA is switched on after every second run of the CA (the "anticipatory" run is figured in), and that approximately 12 rules are allowed to be varied. These are selected at random; the others are blocked for the GA, that is to say, are kept constant. The other systems' values are to be understood accordingly: in system 2, h = 0.5 and r = 0.5; in system 3, h = 1 and r = 1; and in system 4, h = 0.17 and r = 1.

As was to be expected, system 3 wins through: very soon it fills a large part of the screen, and system 2 usually vanishes quite quickly. System 4 can keep up fairly well for some time, but our experiments show that it does not succeed for too long either; system 1 has no chance in the medium-term, either. In view of its size, the differentiation of the predominant system 3 is relatively poor. In comparison, system 1 is much more differentiated. Mathematically, this can be explained by the fact that from the beginning of the predominance of system 3 on, the latter's own system value set the evolutionary standards as it were, so that a maintenance of its own expansionist strategy with a small degree of differentiation was always sufficient for that system. In fact, system 3 became perpetually predominant, specifically by virtue of its permanent use of the GA. The GA of system 3 "learned" very quickly that expansion, and the resulting continual increase of its "population", produced better system values than social differentiation. The other systems, whose "expansive way out" was blocked by system 3, were forced to try to increase their system value by means of internal differentiation. They thereby fell into a trap, however, because with that strategy either no or only a few warriors were generated (since the latter lower the system value for the short term). Consequently, these systems always became a rather defenseless victim of the expansive system 3 in the medium term.

---

[94] This also seems to be realistic: allegedly, Grand Masters of chess rarely plan more than three or four moves ahead because very soon the opponent's actions become unpredictable. What is more important for their success is rather a great repertoire of ad hoc reactions to changing situations of a game, that is, a high variability (cf. Dreyfus and Dreyfus 1987).

There are certainly numerous examples from history where exactly that happened, namely the conquest of evolutionarily advanced societies by more primitive but aggressive ones: Teutons and Rome, Mongolians and China, etc. What is important here, however, is the fact that high variability seems to be the deciding factor when all other starting chances are alike. In general, the relatively cooperative systems 1, 2, and 4, which often established cooperative relations between each other, where not variable enough to meet the aggressive strategy of system 3 with success - they only had their short-term advantage "in mind". Although the above defined values of the pay-off matrix did not change in SOCVAR, system 3 seems to be able to perform something like a variation of the pay-off matrix because of its high variability: it confronted the other systems with situations of decision in which all values of the matrix had in fact become irrelevant except for the aggressor's advantage over non-aggressive systems. These results from SOCVAR show, like the above considerations on possible matrix variations, that game theoretical analyses must consider the problem of the variability of strategic actors even more intensively.

In section 2.3., a strategic player was classified to level 0 according to the learning levels of Bateson. This classification is fairly plausible with regard to the fact that a strategic player acts strictly according to the optimization criteria of the game theoretical possibilities: given a particular situation of decision with a specific pay-off matrix, a player will always use the same strategy - like, e.g., the programs in Axelrod's tournaments. In SOCVAR this is obviously different, and this is so mainly because of the basic moving target situation: a system must constantly decide whether to follow short-term or long-term strategies, and in the case of a long-term decision it is very likely that even if the anticipation size is increased, the system will not be able to tell what is the best strategy. This is not to say that a system can never recognize the *objectively best* strategy, if there is one at all, but that *from the system's point of view* it is not perfectly clear what it should do. Once again, the only thing the systems can do is use the tactics of muddling through, which we have already referred to several times in connection with the GA. That can be successful or not; the only thing that can guarantee the system's long-term survival is a variability as high as possible. Life punishes those who do not change, to vary Gorbachev's famous phrase. Hence, systems like the ones in SOCVAR are certainly not on learning level 0. Instead, they regularly attain level I, and attain level II with regard to the pursuance of strategies that are unfavorable in the short-term, but that are favorable in the long-term.

On the whole, our experiments with SOCVAR, of which we can only describe the one above in detail, confirm the trend regarding h and r that

had already become apparent in our experiments with simpler prototypes (see above, section 3.3.3.): high variability is almost always *the* evolutionary advantage, that is to say that the system's own optimization speed, and its consequent tendency to be superior to competitors, largely depends on the size of h. Interestingly enough, SOCVAR also showed that while h has to be as high as possible, a considerable increase of r over 0.5 does not bring any improvement worth mentioning to the system's competitive success. The trend of the results reported in section 3.3.3, which were obtained in constant environments, seems to be applicable to variable environments as well (the problem of the distinct relevance of rules is rather irrelevant with regard to systems of the type SOZION).

Translated to everyday situations of decision, this does not necessarily mean that one should permanently change one's own strategies, but rather that the *willingness* to change one's own strategies must always be there. It can be perfectly reasonable to hold on to a particular strategy; however, this must not become a principle. By the way, we think that the correspondence between the trends of such dissimilar systems as SOCVAR and the simple prototypes of section 3.3.3, as well as comparable results by Christina Stoica from the analysis of hybrid interactive networks, strongly indicate that these results are no accidents or computer artifacts: they rather seem to reveal fundamental principles of systems in general.

An obvious objection could be raised in connection with the results of SOCVAR regarding their model's adequacy to reality: the fact that evolutionarily advanced societies, especially those that have attained a higher technological level, usually have developed better weapons which enable them to raise the efficiency of their warriors has not been considered at all. To fix this problem it would be necessary to change the model, e.g., by subjecting the warriors' effectiveness (= probability of winning when encountering other warriors) to the existence of craftsmen and technicians. By doing this, societies which are more highly differentiated, but that are smaller in number, would be given a fair chance. In fact, we shall test this in the near future; however, it is almost certain already that variability remains the deciding factor since, e.g., system 3 gained its predominance at equal starting chances only because of its high variability, which also would enable it to quickly increase its technological efficiency, of course. There is, as it were, no escaping Gorbachev's motto.

Systems like SOCVAR can in principle be extended in any possible way, so that more and more factors can be taken into account. It should have become clear after all this that hybrid CAs offer an almost

unlimited richness of possible models. Whatever it is that characterizes societies "per se" - it is obviously possible to model it formally and analyze it by experiment. We will return to that in the final chapter.

Now what are the consequences of all this for problems like that of Hobbes, and also for more practical problems regarding the possibilities and chances of cooperation? Apparently, we first have to distinguish once again between the biological aspect of the problem as it was described by Maynard Smith, Nowak and May, and also Lindgren, and the sociological aspect, that is to say, between the systems of the natural sciences and those of the cultural and cognitive sciences. The former allow the assumption that there is in fact a constant pay-off matrix according to the classical iterated prisoner's dilemma, which favors the development of cooperative behavior as an evolutionary advantage. Here evolutionary dynamics arise through mutations, e.g. by "copy errors" at the transfer of strategies, like in Lindgren's model (see also the biologically orientated model of Parisi et al. 1995). This might have been similar in social systems at the beginning of socio-cultural evolution, at least in a phase when these systems were still very "natural", that is, when they had a low variability (see below, section 5.3.). However, the history of mankind can also be read as a continual increase of the variability of social systems, measured in parameters like, e.g., κ or μ. Of course, an increase of the variability alone does not necessarily have to entail an advance towards cooperativity - on the contrary, as we have seen. In fact, it would be quite unreasonable to interpret history as a constant increase of peace. What does increase is, rather, the evolutionary efficiency, as can be clearly seen in our experiments, and, on the other hand, the obtainable degree of freedom for the evolution of the single systems. Variable systems and, in particular, those that are willing to learn, can answer Hobbes' question practically by giving actors who are willing to cooperate better chances through new variations of the pay-off matrix; we saw that there is no need for a Leviathan to achieve this; we "only" need a high mutation rate, that is to say, a great readiness of the particular individuals to change. If that is not the case, and the matrix is variable so that aggressive actors can continually reinforce their advantages, then a solution from above - Leviathan - or from outside - NATO in the conflict in Bosnia - really seems to be the only thing left to do.

The classical theory of games' search for "the" appropriate strategy in a particular situation of decision has thereby not become obsolete, since one can assume at least for the short-term that, e.g., the pay-off matrix, or other factors that are constitutive for decisions, will not change considerably. However, a single look at the recurrent attempts to use computer programs to comprehend highly variable systems like the stock

market, and to predict share prices (see Schöneburg et al. 1990 as one example of the numerous attempts) demonstrates all too obviously that the theory of games also cannot escape the fact that social systems must be analyzed as variable systems. Hence, the analyses presented here can also be understood as a plea for taking the theory of games as part of a theory of social adaptive, and thus logically hybrid, systems. The sociological mainstream could profit from such a theory of games, not only in special cases, but also in general.

4.4. THE CHARM OF A DISCRETE GEOMETRY (TOGETHER WITH JÖRN SCHMIDT)

In the natural scientific and sociological systems theories we have dealt with so far, the problem of geometric space has for the most part not been paid much attention. This is rather unproblematic in the case of natural scientific systems theories, where the interaction of elements takes place within the physical space which is known to be a frequently and thoroughly examined and explained subject of physics. One can assume without any undue reduction that natural systems operate within classic local Euclidean space-time; we may confidently leave the limitations of this mathematical structure to the theorists of the very small - quantum theorists - , and those of the very large - cosmologists. Aspects of physical space do of course play a part in natural systems, as we could already see from the simple predator-prey system: in fact, a key advantage of the modeling of social systems with CAs, as against the classical Lotka-Volterra equations, is the fact that in CA models the physical space is explicitly included. However, this problem is not a fundamental one, but instead concerns the completeness of models, as we could already gather from the game theoretical model of Nowak and May.

In social systems the problem of space is a more fundamental one, because here it is not necessarily clear from the start what the use of spatial concepts is supposed to mean - apart from pure imagery. We probably all know intuitively what is meant by the "closeness" of two people, or by the growing "distance" between them. In this case, geometric concepts are obviously applied to the different intensity of feelings, and elucidate them. Likewise, we can understand what can be meant by *social* closeness: two individuals are socially close to each other if immediate interactions between them are possible and take place because of their social roles. This is the case, for instance, when occupational roles are the same, or when they are complementary

couples such as doctor and patient. Since the use of such geometric concepts is usually unambiguous in an elaborated ordinary language, no particular specification is necessary here.

However, when we look at the more precise language of computer modelings and mathematical systems analyses, we will not be able to do without clarification for long. This already results from the methodical principle of the "rules of local interactions". The meaning of "local" is perfectly clear in the case of natural systems, since it refers to the neighboring cells in physical space; an application of these methodical principles to the modeling and analysis of social systems that is more than just metaphorical obviously requires a clarification of what exactly is meant by "local" versus "distant" in the analysis of social systems. Of course, people as occupants of social roles act within physical space too; however, this can hardly be the meaning when using spatial concepts in the context of social problems. There are in fact consequences of social interactions in the physical space like it is demonstrated, e.g., by Schelling's aforementioned analyses of segregation; it is certainly true that in general the population distribution to a city's different residential areas mirrors essential features of the social structure and probably reinforces it. However, the space in which social systems and their elements interact is not physical space, but instead is something more abstract. Let us call it the *space of social interactions*, and let us try to define this term more precisely on the analogy with physical space. This principally requires the clarification of two structural characteristics: first, the *topology and metric* of our space, and second, its *dimensionality*. Only after clarifying these basic features it will be possible to speak of the "spatiality" of social interaction in a more than just metaphorical sense.

One thing can already be stated from the outset, however: since the concept of a *rule* lies at the center of all our considerations - be it in the formal or social sense -, a "geometry" of spaces of social interactions can only be understood in connection with the properties of specific rules: in section 3.1.2 we introduced the distinction between "general" rules and "geometrical" or "topological" ones. We saw that the geometry of a system, measured by the topological parameters, has an important impact on the dynamics of the system. Yet the social meaning of geometry was not explained in detail. The following sections elucidate this meaning.

4.4.1. The Topology and Metric of Social Spaces

When mathematicians speak of spaces, they always have abstract sets in mind which allow the analysis of particular problems of mathematical interest, and which are connected with physical space only by generalization. In this study, for example, we often mentioned state and

phase spaces, and also rule spaces. What was meant by that was always the combination of the characteristics of specific systems in abstract sets, which could then be examined as *independent objects*. For example, the concept of a trajectory only makes sense with regard to a state space. At the same time, it is tacitly presupposed here that this abstract and generally highly dimensional space has the topological and metrical qualities that are known, e.g., from the two-dimensional plane. Accordingly, trajectories can be analyzed as independent objects in that space as is done, e.g., in the "CA atlas" of Wuensche and Lesser (loc. cit.). Correspondingly, the control parameters described in section 3.1 are strictly speaking nothing other than trajectories of *rule systems* in an abstract rule space: the variations of the parameter values define "sequences of states" of the rule systems - and not of the systems themselves! - and thus trajectories in the rule systems' state space. Trajectories of hybrid rule systems are characterized by the variation of the meta parameters κ and μ on an even higher level of abstraction in hybrid rule spaces.

The reason for using the term "space" in this and other contexts is the mapping of certain properties and processes of systems into geometric models. Of course, with regard to real systems, these abstract spaces only "exist" in the constructions; what they have in common with the space that is given to our perception are the above-mentioned characteristics, namely the possibility of using geometric relations for the analysis of abstract connections. In the course of the development of analytic geometry since Descartes, it has become quite clear what a powerful instrument this is.

In more technical terms, a space in the mathematical sense is a set for which a *topology* is defined.[95] In simplified terms, it is a structure that is usually defined according to set theory, and that indicates with regard to the single elements of the set which other elements are related to those elements in a special way. In general, topological relations define things like "unity", "separateness", "neighborhood", etc. - in short, the most general geometric relations. The *topological* equality or equivalence of two geometric objects is present if all of their points are in the same topological relation to each other: a circle is topologically equal to an ellipse, regardless of their respective sizes; on the other hand, the circle is not topologically equal to another circle that has a small circle removed from its center.

[95] Strictly speaking, this should be called a topological space; however, one usually presupposes, as was mentioned above, that the respective space has a topology and a metric.

Hence, mathematical topology, understood as the analysis of topological structures, is concerned with the most elementary geometric relations. and especially with aspects like the "density" or "cohesion" of sets. That is what makes topology one of the fundamental logical principles of differential and integral calculus. Note that topology does not concern the measuring of distances; geometric relations of elements are only classified according to whether the elements are in direct or indirect connection to each other. A measuring of distances is effected by introducing a *metric* in addition to the topology of a set; we already dealt with this in section 3.2, in connection with the proof of convergence for the genetic algorithm, and in 3.1.2, in connection with the d parameter (the mean distance between elements of a BN). Let us now introduce it in more detail:

A set is termed a "metrical space" if for every two elements x and y a real number d(x,y) >= 0 is defined with the following qualities:

(a) $d(x,x) = 0$;
(b) $d(x,y) = d(y,x)$;
(c) $d(x,z) \leq d(x,y) + d(y,z)$ (triangle inequality).

As noted in section 3.2., d can be interpreted a bit more illustratively as the distance between two elements: it is obvious that an element's distance to itself is zero; the distance from x to y equals that from y to x; and finally, the distance between x and z cannot be greater than the sum of the distances from x to any other element y, and from y to z. Thus, the above definition corresponds to what we know as distance from our experience with physical space. A metrical space is always a topological one as well, since every metric "induces" a topology.

CAs can easily be characterized as metrical spaces: this is done by selecting a specific neighborhood, e.g., a Moore neighborhood. Then the distance d between two cells x and y is defined as follows:

(1) $d(x,x) = 0$;
(2) $d(x,y) = 1$, if y belongs to the Moore neighborhood of x;
(3) if y does not belong to the neighborhood of x, d(x,y) equals the number of cells that lie between x and y (this number is obtained by taking the shortest path between x and y, e.g., the diagonal, if x and y are not in the same line or column). Thus, in this case $d(x,y) \geq 2$.

Evidently, conditions (a) and (b) of our metric definition are fulfilled by (1), (2), and (3); given that both y and z belong to the neighborhood of x, condition (c) (the triangle inequality) simply derives from the fact that then $d(x,z) = d(x,y) = d(y,z) = 1$ and $d(x,y) + d(y,z) = 1 + 1 = 2$. The validity of (c) can be proved accordingly for the other cases. Hence, the

mapping of a system into a CA allows the description of the relations between the elements of the mapped system as geometric relations, that is, in this case, metrical relations. Note that we are dealing with a "discrete" metric here, which implies that the distance values can only be integers - 0, 1, etc. By contrast, in the physical space there is a continuous metric. We shall see that it makes sense as well to define the metrics of spaces of social interactions as discrete in most cases; therefore, it is appropriate to speak of a "discrete geometry" in general here.[96]

The introduction of a metric to BNs is done accordingly. For two elements x and y

$d(x,x) = 0$

$d(x,y) = 1$, if x and y are connected directly by a BF; else $d(x,y)$ equals the number of cells that serve as the shortest connection chain between x and y. If there is no such chain, $d(x,y) = n + 1$ with n being the number of units of the BN.

Having introduced the mathematically basic concepts, we are now going to consider which sociological questions can be managed with this terminology (with appropriate computer programs); for this purpose, we will make a short excursion into the analysis of social networks; this research line has already developed a tradition of formal and computer based analyses of social relations, and has also been particularly concerned with the "distance" and "closeness" of relational patterns.

> "Network analysis is concerned with the following issues: (1) the empirical description of networks: to what extent do asymmetrical relations exist in a net? Is the network meshed densely or loosely? Are there local condensations in the total net? [...] Can we find typical grouping patterns of actors? [...] This is the basis for network analysis to (2) reveal the consequences of the net structure for the actors' actions. [...] Finally, (3) network analysis wants to find out the reasons for the emergence, the persistence, and the decline of specific networks." (Schweizer 1989, 4, translated by F.B.)

The problems mentioned by Schweizer under point (1) can be understood as problems of genuinely topological relations, and that is the

[96] In a Moore neighborhood of 8 cells the metric obviously corresponds to what is to be intuitively expected. In other neighborhoods these definitions would not be so easily understandable, since, e.g., in a Neumann neighborhood the cells at the core cell's corners do not belong to the neighborhood, which makes their distance to the core cell $d > 1$. Analogously, in a Moore neighborhood of 24 cells, that is, if we include the next lines and columns but one in the neighborhood, the distance from the core cell to any of the 24 neighborhood cells would be $d = 1$, even to the cells that are "next but one".

way they are dealt with in network analysis. The relatively informal description must not hide the fact that these problems have to be formulated and handled very precisely. "Density" or "looseness" of a network, e.g., are concepts that can be defined mathematically. In network analysis this is usually done by means of graph theory; a graph is simply a set of points with (or without) connecting lines between them, namely the relations between the network actors which are symbolized by the points. Now one can roughly determine the network's density by the density of the respective graph, with the latter ensuing as the number of actually existing connecting lines in proportion to the generally possible ones. Correspondingly, local condensations can be studied by means of analyzing so-called subgraphs and their density, and so forth. Presently, network analysis provides an elaborate apparatus of mathematical techniques and relevant computer programs for the treatment of these and similar problems (cf. Freeman 1989).

Metrical relations can also be formally described in network analysis, which can be illustrated with Milgram's "Small World Experiment" (1967): a "target person" is selected from a larger group and several "starting persons" are told to send a "package" to the target person along chains of persons, each of whom knows the next one; so every person in the chain hands the package on to another person that he knows. The interesting point here is of course the number of "intermediary persons" a starter needs to reach the target person; this number provides information about the "density" and "closeness" of the group as a whole; at the same time, the number of intermediaries per starter can be defined as the distance between starter and target person. This is obviously the same metric as in the CA, namely the shortest path between two persons measured in the number of necessary steps (see also Klemm/Wittek 1989). The distance between two points on a graph is defined correspondingly as the length of the shortest path between them; length is simply the number of lines between two points here.

According to these sketches, it apparently turns out that social networks are easily metrizable by the appropriate graphs, so that it is possible now to model all relevant relations within social network as geometric relations. This can indeed lead to non-trivial results, as can be shown by the concept of cyclic graphs; these are graphs where the paths stem from and return to one starting point. A graph containing a cyclic subgraph that comprises all points of the graph is called a "Hamiltonian graph".

In Hamiltonian graphs, the actual network is not stratified with regard to the relations represented in the graph; no individual in such a network can control particular communication paths, since there are always other paths possible. The case can also be reversed: if there are no cyclic

subgraphs, including the graph itself, then we can assume a stratificatory structure of the network; a graph analysis can then show which individual(s) exercise(s) particular controls.

With these rather simple means we are able to define a metric and its corresponding topology for social spaces. The basic idea here, to emphasize the point once again, is that the directness or indirectness of social relations between individuals - or other actors like, e.g., groups - is used to measure topological neighborhood and metrical distance. If two individuals are in a direct relation to each other, e.g., teacher and pupil, they are adjacent, i.e., connected with each other by a connecting line on the corresponding graph, and their distance is $d = 1$. The responsible school inspector is linked directly with the teacher, at least in principle, and connected with the pupil only via the teacher. Thus, the distance (schools inspector - pupil) is $d = 2$, etc. It should be obvious that such a definition of distance fulfils the three metrical conditions, particularly the triangle inequality: given three individuals x, y, and z, there is always either a more direct connection between x and z than via y, or the connection via y is the shortest between x and z, and is thus longer than the direct connection between x and z.

These means allow a very detailed structural analysis of social networks (see also Freeman 1989). However, the issues that Schweizer notes under (2) and (3) are more problematic, since one cannot draw any conclusions as to how the network originated, how it will possibly change, and what the net structure means with regard to the single actors' actions from a purely structural analysis alone.[97] It seems to us that this is mainly considered as an empirical problem by network analysts, although structuralist ethnologists have got a lot to say about problem (2) of course: e.g., the kinship structure of a tribal society predetermines the behavior of the individuals in a specific way: they can only marry particular persons, etc. (cf. e.g. Bourdieu 1992). Thus, we obviously have to understand the social relations as *rules of action and interaction*, respectively, in order to be able to manage problems (2) and (3). With that, however, we actually leave the theoretical frame of structuralist network analysis, and enter the field of a theory of dynamic systems. Let us illustrate this with a classic example, namely the construction of "sociomatrices" within a group (Moreno 1934; Davis 1967):

Given a group of individuals constituted by "external" conditions - the classic example for that is a school class. Every individual is asked to state which members of the group it likes, which are indifferent to it, and

[97] New studies about the evolution of social networks show indeed that until recently the dynamics of social networks has not been picked out systematically as a central theme (cf. Doreian and Stokman 1997).

which it dislikes. These ternary statements can of course be refined, e.g., by a decimal scale, but experience shows that ternary scales are usually sufficient (Moreno himself used only binary scales for the most part, namely friendship and dislike). The statements can now be combined in a ternary coded sociomatrix, that is to say, a matrix where 1 stands for friendship, 0 for indifference, and 2 for dislike.

The construction of such sociomatrices has become part of the standard repertoire of pedagogues and group therapists for a long time; such a matrix, which is also described as an "adjacent matrix" in network analysis, quickly allows insights into the group structure, especially in larger groups: who are the "stars", who the outsiders, what kind of cliques there are, and so forth.

However, such a sociomatrix is only relevant in practice when the relations that are plotted in the matrix are interpreted as potential interactions, which results in a prognosis both for the single actors and for the group as a whole: actors with mutual feelings of friendship - a 1,1 relation - are likely to interact often with each other; actors with 1,0 relations, i.e., friendly - neutral, will do so less often; actors with 0,0 relations (mutual indifference) will only interact by chance, or when it is not possible for them to interact with other actors; and disliked actors will isolate themselves or be isolated. With such an interpretation the matrix allows an evaluation of the group's dynamics that can then be influenced by the teacher or group therapist - e.g., by a seating plan that brings outsiders and popular pupils together in the class.

Together with Ralph Kier, we have illustrated this with a simple CA that operates on the basis of a sociomatrix specified by the user. The group whose sociomatrix is concerned can also be varied in size - from 5 to 80. The matrix components can be entered manually or generated automatically at random. The CA rules simply imply that a cell (a group member) endeavors to surround itself with group members with whom it has as good relations as possible. In practical terms, this means that a cell checks its neighboring cells with regard to its positive or negative relation to them, and then evaluates its social neighborhood (= neighborhood within the CA) by adding up the relation values. Subsequently, the cell checks whether there is a better neighborhood to be found in the extended Moore neighborhood. If this is the case, the cell moves in that direction, otherwise it remains in its social space, that is, it stays with the same confederates. Hence, the only important factor for the behavior of a cell is its unilateral relation to the other cells; the inverse relations do not matter for the behavior of a cell, but they do matter for the behavior of the neighboring cells. All cells are interested in obtaining as good as possible neighborhoods, and so they in principle

constantly change their places, since positive or negative feelings are known not to be necessarily mutual.

The instructive point about these experiments lies in the following result: regardless of the actually selected starting matrix, the number of group members, and the size of the scaling interval - three, ten, or even more values on the "emotional scale" - the CA always turns into a point attractor after not more than 5 or 6 steps. After a short phase of changing social positions, the whole situation of the group quickly stabilizes in small clusters of members that no longer change. Thus, once the group members' mutual emotional values have been settled, the group has lost its chance to form a complex dynamics; it always runs into a specific attractor.

These results are probably well-known to pedagogues and/or therapists with group-dynamic experience; the pedagogical or therapeutic problem in fact often lies in breaking up such rigid group situations, i.e., in the attempt to achieve a more complex dynamics. However, the fact that there are almost mathematical necessities at work is anything but trivial: if there are clear affections and aversions within a group, and if these are transformed into appropriate rules of interactions, then any group, regardless of its other properties, must necessarily end up in such a dynamics. We already mentioned that this can indeed be regularly observed empirically. If this situation should happen to be unfavorable for the group or for particular individuals, then external changes are often the alternative left, because normally such groups are no adaptive, i.e., variable systems. Therefore one must externally institute changes of the interaction rules, introduction of new members, etc., which can be easily done in our formal system of course. In particular, this makes it quite easy to analyze which kinds of changes are the most effective.[98]

So what were mapped into geometric relations in this case were rather emotional relations, even though these relations make up "the social" of the group. As for the rules of interaction, it was supposed that positive emotional relations induce social closeness, i.e., continual interactions, while negative relations induce social distance. Of course, this works with any other kind of relations as well. The decisive point is that the relations must be interpreted as rules of interaction like in the example above. Only when this is possible, on whatever empirical or theoretical

[98] We have abstained (till now) from putting our program at the disposal of pedagogues and therapists, although it would surely be useful. However, Christina Stoica has already used the program with students in one of her seminars to elucidate their own group situation and dynamics to them. She apparently had a sweeping success because the group members could hardly believe the program's agreement with their own experiences. Hence, such programs can have a very practical and enlightening relevance, which we have also experienced in other didactic contexts.

basis, the tasks mentioned by Schweizer under (2) and (3) can be managed; this particularly applies to the changing of networks that was sketched in the above example. In other words: a pure structural analysis, as was frequently carried out in ethnologically orientated network analysis, often remains "sterile", to borrow a quotation from Rapoport (1986 245), unless one is interested neither in the network's dynamics, nor in the corresponding interactions of the actors. A pure structural analysis may be sufficient if there is reason to believe that the network is not going to change in the foreseeable future; and such an assumption is often correct, as it seems, in ethnological analyses of "simple" societies. In general, however, it seems that such an assumption does not necessarily apply so that the transformation of structural networks into the rule systematics of dynamic systems is required.

With that we can also make a concluding terminological clarification regarding the concept of structure now. We proposed in this study to define the structure of a system as the set of rules that generates the system dynamics. The fruitfulness of this definition has hopefully become clear by now. However, if a system is modeled as a network - e.g., by the use of a sociomatrix - then, on the one hand, interaction rules must be specified to be able to grasp the dynamics of the system; and on the other hand, the relations only give information about which specific interactions between the actors are *actually possible*. In other words: the structural analyses, based on specifications of relations, which are mathematically defined in network analysis, describe nothing other than the geometrical rules of interaction, introduced in section 3.1.2. As the concept of structure is usually identified with relations, we also speak, as mentioned above, about the structure of social systems only in connection with the geometrical rules. But it must not be overlooked, as structurally oriented social scientists often do, that the geometrical rules are only one part of the rules of a system and that the dynamics of a system always depends on both kinds of rules (and their respective control parameters).

Regardless of these qualifying remarks, the studies especially in the context of network analysis show how fruitful the geometric modeling of systems is, or, to return to the terminology introduced at the beginning of this chapter, how important a geometry of spaces of social interactions is. Concepts like "distance" or "closeness" define a topology and metric of the social spaces in question quite immediately in the outlined way; this makes it possible to describe social relations in a precise fashion, and to describe them as part of the dynamics of a system. Consequently, this is a first step towards a "geometry of sociology".

Let us return to the topological v parameter of section 3.1.2 once again: it comes out as a degree of the evenness of interactional relations

in Boolean nets. Since the rules of interaction are perfectly clear in BNs, in that they are particular Boolean functions, the *generation of specific dynamics* ensues from particular topological properties of the net in this case, and thus from its geometry. This and the d parameter emphasize once again the relevance of geometric analyses of social spaces: apparently, the latter's geometry can be as important a factor for the dynamics of social systems as the rules themselves. In a way, this makes one think of the famous dream of Einstein, who is known for his aim of putting physics down to geometry. Anyway, in a much more modest way it seems fairly reasonable to put social dynamics down to general rules *and* geometry.[99]

If the elements of social space are roles, then the role-specific rules of interaction obviously determine which roles are close to each other, and which are not. If, on the other hand, the elements are individuals like in the group example above, or "collective actors" like groups or institutions, then these definitions apply accordingly, namely with regard to the degree of directness of the interactions. For example, the distance between sovereign states is $d = 1$; in comparison, a Federal State of the Federal Republic of Germany and France have the distance $d = 2$, because generally a Federal State can only have legally binding interactions with another nation via the Federal Government. Two enemies in a group usually have the distance $d \geq 2$, since a mediator is required in most cases for the two to communicate. In hierarchical organizations the distance values can be even more easily determined, because they are already expressed by the organization structure: which superiors are the subordinates allowed to talk to officially, and with whom may they only correspond indirectly? Here the distance values can be used as a direct measure of hierarchies.

The mentioned experiments by Milgram already demonstrated that distance measuring also makes sense in non-hierarchical groups: obviously, the social coherence in a group is greater the more individuals are relatively close to each other, i.e., the more group members have mutual distance values of $d \leq 2$. Without going into details, we would like to refer to the fact that the theorems of the coherence of social groups as they have been known since Simmel can also be operationalized and stated more precisely in this terminology (for a specification of Simmel's concepts see Sylvan/Glassner 1987). We will

[99] In the meantime, that is, after finishing the German version, we got some more results about the topology of social systems, especially about the influence of particular features of symmetry on the dynamics of social systems. As each book has to be finished some time - even this one - we do not include the new results here; they will be published in a special paper in JASS (Journal for Artificial Societies and Simulation).

be concerned with further problems of cohesion in the following subchapter.

After all, the subject of social spaces gets rid of all its imagery, and can be applied to the precise terminology of a mathematical logical basis, as is usual in this study. The already mentioned discreteness of the metric introduced here is no limitation to generality, in our opinion, because all mathematical characteristics of metrical spaces can also be fulfilled in the discrete. Maybe emotional relations can only be subjectively described as being continual - a little more or less affection; however, this is no problem as regards the techniques of research, since every scaling can be refined in any desired way. Empirically, however - see above - , simple discrete degrees are perfectly sufficient. With that, we also have the possibility of mapping spaces of social interactions with the appropriate metric in a formally adequate way into formal systems like the CA or BN, where we can subsequently analyze the system dynamics, also with respect to its geometric meaning.

From the point of view of model theory, the important point seems to be that topological and metrical relations - relations which are defined as geometric, to put it more precisely - are mapped by a system into the respective model in such a way that these relations are maintained by the mappings. Thus, they are "topology maintaining mappings". The brain presumably works in a quite similar way: specific relations of the perceived objects are processed in the brain - "perceived" - by the process of their being mapped into topological patterns of neurons; these form geometric perceptual clusters, as it were, in which the meaning of the perceptions is constituted for the brain (Edelman 1992; Roth 1996). With a bit of an exaggeration one might say: "meaning" for the brain arises from geometry and from the corresponding rules of neural interactions. In AI research this principle is used in connection with so-called "self-organizing maps": here artificial neural nets map, e.g., logical relations between different sets of information in a way that maintains topological relations into relations between their units (Ritter/Kohonen 1990). What we have here is probably a very fundamental principle of model building, and of the constitution of meaning.

Relations of the kind pertaining to network analysis are of course not the only possible geometric representations of systems. Let us look at two examples at the end of this subchapter to show how diversely social proximity can be modeled as geometric proximity.

(a) In a variant of the iterated prisoner's dilemma, Hegselmann (1996) showed that "solidarity networks" can develop if one merely starts from the egotistical and rational behavior of individual actors, as is usual in the theory of games. In variations on the standard dilemma Hegselmann

introduces actors of different "strength", i.e., actors who regularly win in interactions with weaker opponents, and who hardly have a chance in interactions with stronger ones, respectively. The modeling is done with a CA on whose grid the actors can move. The rules of motion are laid down in such a way that the players can assess their odds with regard to the players in their neighborhood, and can go to a different neighborhood if the adjacent players should not promise success.

In the initial state the players are scattered at random across the grid, and in the resulting final state clusters of players have formed that are more or less comparable with regard to their strength. Hegselmann calls these clusters solidarity networks, i.e., groups in which actors can optimally pursue their common interests and strengths. So obviously, the social closeness of actors that are *similar* in playing terms is represented in the model as geometric closeness; in terms of model theory this means, provided the model turns out to be valid, that the space of social interactions (in terms of the theory of games) is understood as a topological space. By the way, Hegselmann's CA belongs to Wolfram class I, because his evolved solidarity networks do not change anymore, and thus this CA has an attractor of the period 0.[100]

(b) We developed a CA for didactic purposes, inspired by Schelling's analyses of segregation (1971) and by modelings of opinion diffusions (Nowak/Lewenstein 1996) - both using CAs - , to illustrate the development of particular individual attitudes to majority opinions, and the resultant formation of opinion groups. This program, called OPINIO, comprises nine cell states, each of which corresponds to one distinct opinion - e.g., a political one - on a scale of nine values; one can imagine this as a spectrum from the radically left, through moderate values, to the radically right. Basically, the rules say that a cell forms its "opinion" in dependence on its environment, either by conforming to its environment's majority opinion (to put it more precisely: by changing its opinion according to a specific value in the direction of the majority opinion) or by "emigrating", i.e., by changing its social milieu, and looking for one that corresponds to its own opinion. Furthermore, it can change its opinion spontaneously (mutation rate). The corresponding parameters, like the probability of the milieu change, the mutation rate, or the degree and probability of the conforming of opinions by the neighborhood, have to be selected by the user.

[100] It should be mentioned that the term "solidarity" is used quite unconventionally by Hegselmann, since his experiments only consider amalgamations of more or less equally strong players and thus particularly amalgamations of the strong against the weak. It would be better to speak of "alliance networks" here.

With a mutation rate of m = 0 OPINIO usually turns into an attractor of the period 0; with m > 0 it does not. It is characteristic of the program that the opinions become alike at medium values, that is to say, the majority opinions are quite moderate with regard to the whole spectrum. The mathematical reason for this political preference of the center lies simply in the fact that the probability of meeting with moderate opinions in one's own neighborhood is just higher than that of meeting with radical opinions, and thus the success of a totally moderate opinion is actually predetermined - even if there is a predominance of radical positions at the start. At the same time, there usually emerge "opinion clusters", i.e., cells with similar opinions joined together. Interpreting these geometric clusters sociologically, we have a formation of social "grand milieus", that is to say, the close connection and interaction of individuals with almost the same moderate attitudes. Hence, "closeness" and "distance" in social spaces here means direct interaction of individuals with about the same opinions, and only indirect interactions between supporters of very different opinions.

We leave the assessment of the validity of this little model to the reader. Undoubtedly, it parallels aspects of real processes of opinion formation; at the same time, by varying the above parameters it is possible to study the conditions by which there is no occurrence of grand milieus, but the individuals instead remain in small groups, or even isolated. This is the case, e.g., at a high mutation rate where frequent swings of opinion occur - for whatever reasons. The increasing number of floating voters that can be observed in the Federal Republic of Germany during recent years, and the tendency toward decline of the traditional grand milieus can be related to these parameter values. At the same time, a greater independence on the actual neighborhood, i.e., a small probability of alterations and changes of the milieu, obviously impairs the formation of moderate grand milieus.

Further examples could be added, but we can already see from the ones presented here that the concept of social closeness can take many different meanings from a sociological point of view. With regard to the modeling of singular systems this is not really a problem, because the actually intended meaning can be specified sufficiently. Theoretically, however, this is rather unsatisfactory, because it immediately evokes the question of whether there are relations and interactions that are especially important with regard to the society as a whole, and that predetermine a general frame for the geometric structure of the spaces of social interaction as it were. In the mathematical language of vector spaces, this corresponds to the question of a base that "spreads out" the entire space. We think that this question can be answered at least provisionally along the lines of systems theory with the aid of classical differentiation theory,

that is, the theory of socio-cultural evolution within the framework of the different levels of differentiation. In order to give that theory a mathematical, or, strictly speaking, a geometric denotation, it is necessary to turn to our second key word, namely the problem of the dimensionality of spaces of social interactions.

4.4.2. Spatial Dimensions and Social Differentiations

Apparently, the use of spatial concepts in social science contexts in a more than just metaphorical way is not a matter of course. This is especially true for the concept of dimension, which is used to illustrate quite miscellaneous phenomena; however, it usually remains in the metaphorical dark what exactly is meant by a "dimension."

Evidently, the dimensionality of the physical space originates in the experience of three distinct oppositions: in front - behind, left - right, and above - below. It is important here that none of these three oppositions can be reduced to one or both of the others: positions or motions of objects within our three-dimensional experience space can only be fully described as a combination of *three* values which refer to the three oppositions. Thus, the fundamental spatial details concerning individual perceptions always refer to a three-dimensional coordinate system that has *me* as its center; the details themselves are verbalized as topological relations with reference to the center, or as metrical relations - so and so far to the front, to the right, or to the top - in the form of triples.[101]

Being generalizations of the physical perceptual space, the mathematical definitions of the dimensionality of spaces start from the basic experience of their irreducibility. We abstain from giving the fairly complex topological definitions of dimensionality here, and start instead from the (logically derivable) concept of dimensionality for vector spaces, which is perfectly sufficient for our purposes. According to that, the number of a vector space's dimensions derives from the *number of linearly independent vectors* that are sufficient to "spread out" the entire vector space, i.e., to describe every single vector as a linear combination of the units' vectors. The meaning of this definition actually corresponds exactly with the description of objects and/or events in the physical perceptual space: if exactly n different details are necessary to give a complete spatial description of an object/event - no more and no less -

[101] It is known that since the special theory of relativity in physics one usually speaks of the four-dimensional space-time, which is actually true for perceptions as well: something was (before me, above me, to my left, yesterday), so that complete details should actually always be given as four-dimensional vectors. Of course, the time as the fourth dimension can not be reduced *a fortiori* to one of the three spatial dimensions either. However, this will not be pondered any further in the following.

then the number of dimensions of that space is n, in other words: that space is n-dimensional. The four-dimensional space-time is four-dimensional because any complete description must consist of four coordinate details; the much higher dimensionality of the Hilbert spaces in quantum theory results from the impossibility of describing them in a simpler way.

When we introduced the concept of "state space" we used exactly this definition without further explaining that: since our repeatedly mentioned fox and goose system comprises a total of 54 different cell states, which cannot be reduced to each other, the state space has to be 54-dimensional; it is only "pressed together" to a two-dimensional space for the sake of lucidity. To summarize our remarks, we can say that, formally speaking, the dimensions of a space are those characteristics of the space set that cannot be logically expressed either by the other characteristics or by combinations of them.

So much for the basic concepts. The number of dimensions of a physical space is obviously not arbitrary, although mathematicians are long since used to calculating with spaces of any number of dimensions. For example, it would have been hardly possible for complex creatures to develop in a two-dimensional space. One reason is that creatures in a two-dimensional space could only possess one single opening for the ingestion and excretion of food, otherwise the two-dimensional being would be separated into two pieces and fall apart. Likewise, a closed circulation of blood would split the creature into an inner and an outer half, which would also fall apart. Thus, life in a two-dimensional space would be reduced to the complexity level of single-celled organisms (cf. e.g. Hawking 1988, 206).

On the other hand, in a space with four or more dimensions, the laws of nature as we know them would lead, among other things, to a high instability of planetary systems and the dynamic equilibrium of fixed stars, which would practically exclude the possibility of the evolution of life for the long-term. Consequentially, complex creatures, and especially human beings, could only have developed in a three-dimensional space. Hence, with Kant we can say that the three-dimensionality of the physical space is a condition of the possibility of the evolution of complexity (though in referring to space Kant meant this in a different way of course).

After this little excursion into the dimensionality of mathematical and physical spaces, let us now return to our initial problem: it is surely possible to speak of the dimensionality of social spaces in very different ways using the precise terminology of mathematics. For a definition regarding the "society as a whole", however, as was already mentioned, the classical (three-level) differentiation theory suggests itself, which has

become a part of the standard repertoire of theoretical sociology, and is used by rather different authors as a matter of course (e.g. Mayntz 1988; Luhmann 1984; Habermas 1981). The differentiation theory is known to deal with systems rather then individual actors or social roles; nevertheless, the basic statements of differentiation theory can be quite easily transformed into interactionist statements:[102]

A segmentarily differentiated society corresponds to a one-dimensional interactional orientation, in that its interactions are basically orientated to the binary code "familiar - strange". Rules of social interactions *prima facie* only apply to members of one's own society; strangers are either not human beings, or else they have to be integrated into the society by means of additional rules which are in most cases valid only for these special interactions. Numerous ethnological analyses of the rules of interactions between different tribes, e.g., the famous Kula chain (Malinowski 1944), or the institutions of inter-societal marriages (Habermas 1981) clearly point to that. Thus, what we have here is a parallel to the "left - right" orientation.

The transition into stratificatory societies adds a new binary coding, namely the social "above - below". This has been described so often (e.g. Habermas 1981) that we merely have to note it here. The coding "familiar - strange" does not disappear of course and so a second dimension evolves, in fact a form of social orientation that is independent of the first, and that cannot be reduced to it. One reason for that is that this second form emerges within the familiar society. The history of class struggles, from the quarrels between plebeians and patricians in ancient Rome to those of industrial capitalism, demonstrate what price had to be paid for this new evolutionary level.

In conclusion, functional differentiation structures social interactions according to the coding "active - passive" and "action role vs. client role". This results from the "universal inclusion" of the functional systems which has been frequently described in particular by Luhmann: every member of a functionally differentiated society has the in principle possibility, and the formal right, to participate in the achievements of individual subsystems by taking on a client role, even though, in comparison, the adoption of an action role is tied to numerous formal and material prerequisites. Hence, interactions constituted by functional differentiation orient to the fact that the actors meet as (professionally) acting persons and clients, along with special cases of two competing or

[102] One of the few exceptions we see in this context is the well-known model of Eder (1976) and Habermas (1981) where the transition from segmentary differentiation into stratificatory differentiation is described as the transformation of single social roles into new roles.

cooperating acting persons and/or clients. Since actors who interact as occupants of two complementary roles - teacher vs. learner, doctor vs. patient, producer vs. consumer - are *logically* independent of the orientations "above - below" and "familiar - strange", this is another new dimension of social spaces, namely the third one.[103] The specific rules that these interactions follow are logically and sociologically independent of those of stratificatory and segmentary differentiations, at least in principle. After all, this is the logical mathematical meaning of the self-dynamics of (instrumental) rationality and its detachment from stocks of tradition that has been dealt with ever since Max Weber.

It is an empirical question to what extent the functional differentiation and the resultant rules of interaction have become the predominant forms in modern societies. Both Habermas and Luhmann are known to assume that. Anyway, the emergence of functional differentiation certainly did not make the traditional forms disappear (Hondrich 1987), so that in principle we can speak of a three-dimensionality of the spaces of social interactions with respect to modern societies. The extent to which modern societies "unfold" a third dimension in addition to the two-dimensionality of premodern advanced civilizations corresponds to the degree to which functionally constituted rules of interaction turn up as logically independent rules beside those of traditional differentiation.

Before specifying and extending these general considerations with a formal model, we are going to clear up in general what is actually meant by the dimensionality of spaces of social interactions. In spite of our initial remarks about the concept of dimension in the physical space, it is important that in the case of social spaces we are only dealing with a *mathematical* meaning. The only, yet sufficient, justification for speaking of topology, metric, and dimensionality in connection with physical as well as social spaces is that it is possible and necessary in both cases to map the respective rules of interaction into mathematical structures which are well-known and have been analyzed as metrical and multi-dimensional spaces. Social interactions take place "within" a social three-dimensional space only insofar as the characteristic rules of these interactions can be described by the geometry of multi-dimensional spaces. This is not only possible, it also makes sense in terms of sociology, which follows from the above sketches about the irreducible

[103] Interactions that are so to speak functionally constituted are of course not totally free of the first two orientations. Complementary relations like that between teacher and pupil or doctor and patient almost always need a minimum of mutual confidence. Correspondingly, the professionally acting person in such situations is usually, though not always, in a socially higher position than the client. However, this does not constitute the specific quality of those interactions but it just points to the fact that empirically there are no logically "pure" forms of interactions. In comparison, see, e.g., McCarthy 1986 who deals with this in a different context.

orientations of the differently constituted interactions; however, the "space" of social interactions only exists as such because the respective social systems have developed rule systems in the course of their evolution which show geometric characteristics in the *mathematical* sense, especially with regard to the levels of social differentiation. The only reason why it is legitimate to speak of the dimensionality of social spaces, and of social spaces at all, in non-metaphorical terms is that it is impossible to reduce functionally constituted rules of interaction to stratificatorily or segmentarily constituted ones, in other words: to describe the former as a linear combination of the latter two.

This is, *mutatis mutandis*, also true for the topology and metric of social spaces. Just like there is the fundamental social experience of orientations, which makes up the basis of social differentiation, there is of course the fundamental experience of the socially "close" or "distant"; correspondingly, any normally socialized member of a society knows the difference between "dense" or "coherent" groups and rather "loosely coherent" groups. We could see that, and to what extent, the topology and metric of social spaces can be derived from these basic experiences; this is necessary, however, to be able to use the concept of social space precisely.

Nevertheless, it is at least of heuristic use to recall the *physical* meaning of the space concept from time to time. We noticed above that a third dimension of the social space "unfolds" in modern societies. Likewise, we could say that a second dimension unfolds at the transition from segmentarily differentiated societies into stratificatory advanced civilizations. This is to say that social systems are not three- or two-dimensional per se, but that the dimensionality of their spaces of interactions is a result of their evolution. By developing, varying, and adapting their rules to new environmental requirements as adaptive systems, social systems also develop the geometry of their spaces. This obviously stands in contradiction to the common view of the relation between space and geometric structure on the one hand, and the evolutionary processes "within" this space and its geometric structure on the other - in spite? of Kant's transcendental philosophical reflections, and his followers who were occupied with the problem of space.

However, physicists since Einstein are used to describing the local metric of the physical space as a consequence of the distribution of matter, and thus as something that depends on the (local) evolution of the universe. Present unifying theories even go one step further in that they assume that "actual" space-time is not four-dimensional but 10- or 26-dimensional. According to that theory, four of these 10 or 26 dimensions have unfolded in the early stages of the universe, while the others

remained rolled up within the subatomic small - they were only relevant for the initial states, so to speak (cf. e.g. Peat 1988).

Such ideas may be just mathematical speculations. It is noteworthy in our context, however, that the assumption of complex systems being able to bring about their own geometry, and also the dimensionality of their spaces by their evolution, is obviously not paradoxical at all. The physical conception of rolled up dimensions, four of which have unfolded, can be applied to social systems: the orientations of "above - below" and "actor - client" were of course already there in purely segmentarily differentiated societies. However, they were not constitutive for the binding interaction rules of the society as a whole; they instead remained rolled up in the sense that they only became relevant in local ad hoc interactions. Hence, the chief of a tribe was socially above the ordinary hunter; this only applied to specific interactions though, and did not constitute the society as a whole. Thus, socio-cultural evolution also means the unfolding of additional dimensions.

For a reconstruction and a detailed analysis of all this with a formal model, the use of the geometry of CAs as a formal mapping suggests itself, as in our previous examples. Topological-metrical concepts such as "closeness" and "distance" immediately ensue from the CA's grid construction, which was shown above. However, it is not very useful to represent social dimensions directly by the dimensions of a CA for several reasons, which are insignificant here. Instead, we decided on using the differentiating social dimensions "only" analytically, and displaying the processes of the CA on a two-dimensional level. This results in the following model:

We start out from the assumption that every cell in the CA, which corresponds to a particular individual, is determined by three different variable index values. This means that every cell can be in three different states *simultaneously*, with the first state corresponding to the membership of a particular family (= social segment); the second corresponding to the membership of a particular social class; and the third state corresponding to the membership of a particular functional system; for the sake of simplicity, only the action roles are taken into account. Now the relevant rules of the CA determine (a) how an individual can change from one segment into another; (b) how socially upward and downward mobility from one class to another takes place; and (c) how the change from one functional system to another can be achieved. One can conceive of this program, which bears the pretty name TRISOC (three-dimensional society), as a three-dimensional variant of

SOZGEN - which is one-dimensional in this respect, since TRISOC also has rules concerning the dynamics of population.[104]

The evaluation function is of special importance in TRISOC, because the program is of course hybrid, that is to say, it is controlled by a GA. As regards the distinct social dimensions, the function is basically similar to the evaluation functions of SOZION and SOZGEN, which are three times applied here. In addition, there is a "cohesion factor": this means that a close cohesion on the segmentary level, the class level, and the level of the functional systems are rewarded, respectively: the geometric closeness in the CA is interpreted as social "density" of the single social levels, and is regarded as positive for the system as whole. Behind this is an assumption that has been known at least since Simmel, namely that social density can be regarded as strong integration of a group, and that it is a positive value for that group. Thus, it is the job of the system as a whole to arrange its own evolution in such a way that all three social dimensions can *simultaneously* attain values as high as possible; additionally, it has to keep an eye on the cohesion factor. Furthermore, TRISOC cannot only vary the probability parameters of the rules and switch them on or off, it can also vary antecedens conditions of several rules; making, e.g., transitions between classes big or small (it can decide on whether transitions are only possible from one class to the next, or also to the one after the next, or to the one after that); and it can change environmental conditions. As a result, the whole program is as variable as one could possibly wish for from a logical point of view.

The basic theoretical assumption here, as was already indicated at the end of the last subchapter, is that the social relations which are constitutive for the *society as a whole* must be located along the three axes which correspond to the three dimensions of differentiation in the CA. As we could see in the last subchapter, there are extremely many different social relations, parts of which are independent of the differentiation structure of the society as a whole. But in so far as social relations are directly constituted by the total social structure they can be either associated to one of the three dimensions or described as a combination of relations, each of which refers to one of the three dimensions in mathematical terms. For example, the social relation of an industrial worker to his superiors can be described, first, as a vertical relation, namely as a subordinate one; second, as a mutual membership of the same company, against which other companies are "strange"; and

[104] The processes of SOZGEN or SOZION are of course displayed in a two-dimensional CA, or, to put it more precisely, in a two-dimensional grid. However, interactions only take place with regard to the class structure; therefore, sociologically speaking, SOZGEN is one-dimensional.

third, as the equally mutual possession of action roles in the functional system "economy". Thus, the roles "industrial worker" and, e.g., "factory manager" could be described in a formal way so that (a) both roles belong to the same or a different segment - a company; (b) they are vertically above or below each other; (c) they both belong to the same functional system. With this three-dimensional "spatialization" we can describe the geometric relations between any social roles accordingly.

Like SOZION or SOZGEN, TRISOC normally starts with the "primordial state" of a purely agrarian society that is only segmentarily differentiated. If specific environmental requirements make it necessary, TRISOC introduces the second level of differentiation by launching the stratificatory differentiation in addition to the segmentary one. If this is also insufficient, the third level of functional differentiation is initiated. Segmentary differentiation in TRISOC always entails the division into different family or kin groups; of course, it is also possible to choose any other segmentary differentiation that is important to the society as a whole.

Graphical illustrations of TRISOC show the single social levels on top of each other, with the lowest one being the segmentary level, above that the stratificatory one, and the functional one on top. Social evolution is modeled according to the "density" of the single levels: if all the cells that belong together are tightly connected, the level is "dense", i.e., the society shows a strong cohesion, and thus a strong inner integration on this level. Correspondingly, frayed clusters indicate that the social integration is more and more at risk.

Let us report some of the most important results from our experiments with TRISOC; in fact, we are constantly surprised at new discoveries, and the results may presently only be seen as being rather preliminary. It takes rather long series of experiments to comprehend adequately a program as complex as TRISOC:

(a) Only if the parameters are very carefully chosen - particularly the contribution values for the single social units or the cohesion factors - can it be achieved that TRISOC exits the state of purely segmentary differentiation at all. The maintenance of the simple level of differentiation is mathematically most probable. Only if the beginning of the next differentiation level brings an *immediate* systemic advantage for TRISOC, is the new level of differentiation continued; in any other case it either lingers with a rudimentary accomplishment of stratificatory differentiation, or even regresses to purely segmentary differentiation again.

This indicates something that has been frequently described, though not explained, by sociological evolutionary theorists (e.g. Habermas 1981 I): the transition from segmentarily into stratificatorily differentiated

societies has happened very rarely in full degree. We know numerous segmentarily differentiated societies that became stuck on a level between purely segmentary differentiation and fully accomplished stratificatory differentiation as it were; most of the known societies did not even get that far. In comparison, the genesis of stratificatorily differentiated societies is a phenomenon that did not happen more than ten to fifteen times among hundreds, even thousands of societies. Our experiments with TRISOC can offer some primary possibilities for a mathematical explanation here: the process itself depends on so many different variable parameters that the realization of stratificatory societies seems to be mathematically rather improbable. It requires at once extremely severe environmental problems, as well as the above-mentioned favorable parameter combinations, to cause the generation of advanced civilizations. Such systems can of course only be understood as stochastic systems, and thus the coincidental occurrence of both conditions is obviously very improbable. We will study this phenomenon in more detail, but the trend is already apparent.

(b) TRISOC reaches trajectories that show at least the beginnings of "order", i.e., simple attractors, only with great difficulty. In particular, if the third social dimension (functional differentiation) has been reached, TRISOC realizes at best partial trajectories, i.e., trajectories that only take singular classes, families, or functional systems into account, which yield attractors in relation to each other or to the environmental vectors. Aggregating some or all of the social units to greater parts, or analyzing overall trajectories - which can of course no longer be charted - results in trajectories "at the edge of chaos" (Langton 1992; Packard 1988), or even in strange attractors.

The meaning of these results for social systems of such complexity is that they must continually adjust their own dynamics, that is to say, they permanently have to exert an influence on the particular base rules through their meta rules. No system of that kind, that is, no modern society, in the usual sociological sense, can rely on the fact that everything is going to be alright "somehow": they instead have to prepare themselves for the mathematical certainty that without permanent corrections their circumstances (may) very soon become chaotic. Since it is evident that modern societies, in contrast to SOZGEN, do not have the possibility of solving this problem by simply switching off as many rules as possible after having attained favorable states, the only thing that these systems can do is the compulsion of constant variability. The imprudence of pure laissez-faire, which was already noted by Bateson, is strikingly confirmed here (see below, section 5.3., and the coercion to reform in 3.1.2).

Another fact to be mentioned in this context is that TRISOC is extremely sensitive to almost any kind of external variation of parameters, i.e., parameters that TRISOC itself cannot influence in the process. We have already mentioned in section 2.1 that such a sensitivity is a characteristic feature of chaotic systems and systems at the edge of chaos. Modern societies seem to be systems which are susceptible to interference; only if they can vary practically all of their relevant parameters themselves, can they escape from the edge of chaos far enough so that they are not constantly endangered by influences of their environment. In other words: the stability of modern societies lies in their variability.

(c) Finally, our results show quite clearly that it is obviously too much for TRISOC to accomplish the different levels of differentiation, and at the same time keep up its social integration on all three levels. The levels usually split up, especially those of the segments, as soon as the next differentiation level is introduced. In fact, TRISOC always has the following problem: either the cohesion factor is set to a relatively high degree; then the system is primarily orientated to social integration. In that case, the evolution of the different levels of differentiation can almost certainly never occur, because TRISOC can see no immediate advantage in that - see above. Or, on the other hand, the cohesion factor is reduced so that a system evolution takes place. Then TRISOC is forced to sacrifice social integration on the lower levels in favor of the higher total system performances. Apparently, it is impossible to achieve both - optimally efficient rules and socially dense integration - at the same time.

The parallels of these results to the frequently noted problems of integration and splitting up in modern societies are so evident that we do not need to comment on them any further. Our experiments teach that three-dimensional differentiations almost necessarily lead to situations where traditional values and familiar contexts fray and are in danger of declining. It is probably no exaggeration to point to the famous passage from the "Communist Manifesto" in this context, where Marx and Engels expressed just this with an amazing far-sightedness at the very beginning of this process:

> "Die Bourgeoisie, wo sie zur Herrschaft gekommen ist, hat alle feudalen, patriarchalischen, idyllischen Verhältnisse zerstört. Sie hat die buntscheckigen Feudalbande, die den Menschen an seinen natürlichen Vorgesetzten knüpften, unbarmherzig zerrissen und kein anderes Band zwischen Mensch und Mensch übriggelassen als das nackte Interesse, als die gefühllose "bare Zahlung". Sie hat die heiligen Schauer der frommen Schwärmerei, der ritterlichen Begeisterung, der spießbürgerlichen Wehmut in dem eiskalten Wasser egoistischer Berechnung ertränkt. [...] Alles Ständische und Stehende verdampft, alles Heilige wird entweiht,

> und die Menschen sind endlich gezwungen, ihre Lebensstellung, ihre gegenseitigen Beziehungen mit nüchternen Augen zu betrachten." (Marx/Engels, Manifest der Kommunistischen Partei, MEW 4,464)[105]

If we abstract from the specific terminology, this quotation reads (for us) like a comment on TRISOC; in particular, such programs can tell us to face such processes neither with resignation nor with a critique of culture, but "with sober senses".

One can learn from our experiments that modern societies, as three-dimensionally differentiated objects, are obviously systems that have paid dearly for their immense efficiency: continual compulsions toward competition, and therewith the obligation for permanent variability; and furthermore, the fundamental problem of the constantly threatened social integration. If in such a situation additional, externally caused problems supervene, such as ecological endangerments, or the development of a world society with extremely different developmental stages of its regional spaces, then it is almost a wonder that on the whole the modern societies seem to get along quite well. This probably works only *because* these societies have willy-nilly made the principle of high variability a part of their social identity. A lot of research remains to be done on that.

At the end of these considerations on dimensionality the question of whether this process can (and will) go on in the same way, that is to say, if a fourth dimension will ensue, almost inevitably springs to mind. We cannot even speculate on this question, since presently it is not possible to see what such a dimension would look like. Any present statements about that would merely be science fiction. The only thing that is for certain is that further levels of differentiation would even increase the already noted problems of extreme sensitivity and social integration. It remains to be seen whether modern societies will be willing to pay this price; at any rate, now, in contrast to the stages of the present evolutionary level's development, we can be quite certain about the price that would have to be paid. Even Engels and Marx, despite all their far-sightedness, have understood the development they noticed only as a problem of specifically capitalist societies; experiments like the ones with TRISOC show that this is a price which must be paid by any

[105] "The bourgeoisie, wherever it has got the upper hand, has put an end to all feudal, patriarchal, idyllic relations. It has pitilessly torn asunder the motley feudal ties that bound man to his "natural superiors", and has left no other nexus between man and man than naked self-interest, than callous "cash payment". It has drowned out the most heavenly ecstasies of religious fervor, of chivalrous enthusiasm, of philistine sentimentalism, in the icy water of egotistical calculation. [...] All that is solid melts into air, all that is holy is profaned, and man is at last compelled to face with sober senses his real condition of life and his relations with his kind." (Marx/Engels, Manifesto of the Communist Party...)

modern society, for mathematical reasons. This can not even be helped by a socialism that is better than the caricature of the "real existing socialism".

4.5. SOME THOUGHTS ON THE DEVELOPMENT OF THE NEW

The formal systems described so far, and the problems that were analyzed with them, still have a fundamental deficiency: the adaptive changes performed by the systems mainly consist of *recombinations* of already existing rules or elements, with the exception of the development of dimensions. The operation of the GA is exemplary here: its crossover operator can be regarded as a paradigm of recombinatory changes. However, we could see above that even the mutation parameter can "only" produce changes within a given space of possibilities. We might also say that the innovations of the systems described here involve purely qualitative changes, while the *quantity* - the number of rules and the types of elements - remains constant (though the number of elements may change of course).

According to Bateson (1979), one might say that new ideas have their source mainly (perhaps not completely) in the new distribution and recombination of ideas we already have. In that case, recombination would really be the principal innovative mechanism. Still, there can be no doubt that innovations also entail quantitative extensions of the already existing ideas for complex systems; systems do not only recombine old rules and types of elements to create something new, but they also introduce additional rules and element types which do not just displace the old ones. Let us illustrate this with a classic example from the history of science:

It has been frequently emphasized in connection with the works of Zilsel (1976) that, and how, the modern natural sciences have developed as an "integration" of the social roles of craftsmen, scholars and artists (Böhme et al. 1977; Münch 1988). Hence, the social role of the scientist was constituted from different components of already existing roles; so, again, what we have here is a form of recombination. However, the absolute number of social rules and types of elements increased thereby, and a new social subsystem evolved; the reason for that is that the old roles and their corresponding rules of interaction did in fact not disappear in the process, but coexisted with the new roles and rules. Similar examples could be increased ad libitum. So, according to Bateson we can state that recombination has a central function for the development of the new, since it is obvious that neither social nor natural innovations emerge out of nothing - like Athene from the head of Zeus; consequently, we then have to distinguish between cases where innovatory recombinations

do not yield any quantitative increase, and those where there is a growth in quantity. We may assume that social innovations are open to both potentials in the sense that it is possible for the new to replace the old, at least in the medium term, while at the same time the old can coexist with the new as well. Examples for the first case can be found whenever new forms of rule replace older ones, e.g., in the case of the transition from a segmentarily differentiated tribal society with its temporary rule of the chief, into a stratificatorily differentiated advanced civilization with its permanent king's rule (cf. e.g. Eder 1976).

However, biological as well as socio-cultural processes of evolution mainly - though not exclusively - consist in the fact that the evolving systems increase in quantity. It is obvious that this involves an increase in complexity, both in its algorithmic and in its effective meaning (see above, section 2.6.). A recombination of the kind we just sketched may well also be assumed to play the most important part with regard to biological evolution: the enlargement of a species' genome can in fact only happen when the known components of the DNA are combined in a new way. Let us therefore take a closer look at the concept of innovation as an extension of a system.

It has been known for a long time in chemistry and biology that the basic mechanism of system extensions is that of *autocatalysis*. This means that a (bio-) chemical system can produce new types of elements without the effect of external catalysts. Presumably the most famous example in this context is the hyper cycle of Eigen (Eigen/Schuster 1979), which has already been mentioned. The above example of the genesis of modern science can also be understood in an autocatalytic sense: the social system of the Upper Italian and Dutch trading cities generated the new social role of the modern scientist autocatalytically, namely through the interaction of several already existing roles.

Of course, there is - once again - an important difference between the concept of autocatalysis in the natural sciences and the development of new social roles and rules: whenever new social roles emerge, the social system has to generate new rules which determine the interactions between those new roles themselves and *between the new and the old roles*. In other words: the system has to make sure that the new types of elements can be "joined" to the old types and their rules. This sounds trivial, and in the case of natural scientific autocatalysis it really is, because there the same rules that let new elements emerge autocatalytically from the old ones control how these new elements interact with the old ones and with each other. In fact, they are the constant laws of nature. Things are otherwise in the case of social systems: they can develop various solutions, which can already be seen

from the example of the scientist: the emergence of the social role "scientist" was inextricably linked with the problem of how this new role could be joined to the previous rule system. The declaration of the Royal Society has become famous in this context; one of its essential rules regarding the new roles was that of political neutrality - "not to meddle with politicks" (cf. van den Daele 1977); the institutional integration of the scientist's role was also a continuous problem, spanning from the early scientific academies, through the German 19th century research university, to modern industrial research and onwards (see, e.g., Ben-David 1971).

Therefore, the natural scientific concept of autocatalysis should be used with caution, and only in a metaphorical sense, in social scientific contexts. I rather use the concept of *self-enlargement* in order to be able to include the more complex case of social systems. The two resulting problems for self enlarging social systems are: first, the ability to join the new to the old - irrespective of whether old things are replaced by the new or not, since a system never reshapes itself in a totally new way; and second, the trend to become too large, which entails growing difficulties in accomplishing adaptive solutions to newly occurring problems. Both problems are factually connected.

In the context of the joining of the new, we have to keep in mind that a system usually introduces new types of elements and rules in order to be able to cope with specific problems for which the established rules and element types are not sufficient even if they are recombined. Thus, the system cannot join the new "somehow", but instead has to make sure that this happens in a way which improves its own performance - otherwise it could do without the enlargements. As a consequence, its innovations may not have a disturbing effect on the established rule system, or else the latter has to be changed as well, so that the new rules and elements can take effect. So, again, the system has an optimization problem, which we have become quite familiar with in the meantime; of course, unambiguous solutions can be least expected here, so that the system has to choose the easiest practicable solutions from the space of possibilities of adequate solutions, on the basis of its previous structures. The different manifestations of the scientist's role in the national systems of England, France, and Germany demonstrate that the single systems really carried out the solutions that came closest to their actually established social rule system, that is to say, that could be integrated most easily (cf. Ben-David loc. cit.).

Stoica (1997) has modeled self enlarging systems formally, using hybrid interactive networks. Here, a network of a particular size is optimized by a GA; if the hybrid system does not, or does not sufficiently, obtain the required results according to previously

determined criteria, then the network has the option to enlarge itself; this means that a new unit is added and the weight matrix has to be enlarged accordingly. In the case of n units the matrix consists of n^2 weight values, so that an enlargement of the network by one unit entails that the matrix has to be enlarged by 2n+1 values; so now the hybrid system has to perform a new optimization of a much larger matrix. Stoica's experiments showed that this quickly leads to the effect that, on the one hand, the required optimization values are achieved "in principle" by the network's self-enlargement, but on the other hand, the period of time necessary increases accordingly. In relatively large networks it might be better to do without the enlargement because of the time factor, and to be content with the results of the unenlarged network, which may be suboptimal, but can at least be attained in acceptable time.[106]

One can imagine several solutions to avoid this problem, almost all of which have been accomplished in natural and social reality. One possible solution is to introduce "constraints", i.e., to restrict the effect of optimization and variation processes to only a part of the system. Technically, this corresponds to a determination of the r parameters at $r < 1$. However, in order to rule out unsatisfactory solutions, another determination has to be made, namely the factual separation of subsystems from the system as a whole. Only then is it possible to let the new rules and elements vary independent of the general context. Luhmann has picked this aspect out as a central theme several times, though in a different terminology; one may assume that the differentiation of relatively autonomous subsystems happened exactly for that reason: the entire systems could no longer be optimized sufficiently, as a consequence of their tendency to explode exponentially.

Another solution seems to be accomplished by the brain during ontogenesis in early childhood (Edelman 1992; Müller 1991). The neurons, as well as the corresponding synaptic links, that are existent at the time of birth are drastically decreased in number during the first extra-uterine months; it is estimated that this reduction approaches 40%. So the brain itself reduces its actual unmanageable size by determining during the first months how large it has to be in order to be able to fulfill its tasks. It starts with a surplus capacity - for whatever reason - and subsequently reduces its capacity to the size that is de facto needed.

Social systems are capable of doing that as well. We need only look at some examples from times of crisis: temporarily superfluous rules are suspended and specific rules for managing a state of emergency

[106] We are presently making analogous experiments with self enlarging CAs, which have not been finished yet; the problems are even more serious here because of the systems' combinatory capacity.

predominate almost without restriction. The system's capacity - and therewith its complexity - is considerably reduced. It now depends on the social system's variability whether, and to what extent, its original capacities are restored after the end of the crisis, or whether the system continues to operate with its reduced capacity. It is obvious that a totalitarian system like that of the Nazis or the Stalinist one is by definition determined to achieve permanent, irrevocable reductions of its rules.

So far, we have only talked about *self-enlargements*. Of course, systems can also be enlarged by external interventions: ecosystems increase by immigrations of new species; social systems are forced by superior opponents to introduce particular innovations, like, e.g., Germany after 1945; migrations from other cultures import new rules to which the social system has to adapt itself by innovations on its own part; and so forth. Theoretically, however, these cases do not differ fundamentally from those we were dealing with under the aspect of self-enlargement: the ensuing problems that have to be managed in one way or another are the same here. The system's basic variability seems to be a decisive aspect for the way it deals with the ensuing problems. The problem of migration from foreign cultures, which is highly topical in Western societies, bears a formal parallel to the problem of ecosystems' capacity to include different species that was examined by Pimm (1991), in that it confronts these societies with externally caused problems of extension. It is to be hoped that these problems are not solved by rule reductions, that is, by the renunciation of tolerance and the right of asylum; the highly variable tradition of the Western societies may perhaps give rise to some optimism here.

Stoica's experiments indicate that the meta parameters analyzed in section 3.3.3. can be applied accordingly to the problems of (self) enlargement, and that the present results - with appropriate adjustments - are valid here as well. Still, as we have repeatedly emphasized, this is yet to be proven by further experiments. Nevertheless, and regardless of these open questions of research, our sketchy remarks have probably already demonstrated that, and how, the phenomenon of self-enlargement of social - and other - systems can in fact be managed in a formal way. Social systems achieve self-enlargements, which they can also undo again, on the basis of their adaptive competencies; hence, they cut out those solutions from the space of possibilities that fit *their history as much as possible*. Of course, an exact prognosis is only possible in very simple cases, but this applies, in any case, to any kind of adaptive variation (see below, section 5.2.). It is a question of practicability whether and how innovations are adequately joined, and which additional measures have to be taken. In a somewhat metaphorical

terminology we might speak of an "innovatory complexity" here; in plainer terms, we are simply speaking of another facet of the adaptivity and variability of complex social systems.

5. RULES, UNIVERSALS, AND QUESTIONS OF RESEARCH - A CONCLUSION THAT IS NOT AN ENDING

> "Ich behaupte aber, daß in jeder besonderen Naturlehre nur so viel eigentliche Wissenschaft angetroffen werden könne, als darin Mathematik anzutreffen ist." (Kant)[107]

It was inevitable that the main part of this study dealt with formal systems and their properties while relatively little has been said about sociology, although I attempted throughout to refer to the characteristics of social systems. This "sociological deficit" ensues from the nature of the things we are dealing with here: if sociology is to be conceived as a stricter science than in its theoretical traditions, and if this is to be achieved without sacrificing the essential features of social systems in a reductionist way for the sake of mere exactitude, then it is necessary to set up its fundamental concepts and methods first, and to put them in concrete form by the use of exemplary models. On the other hand, it is equally necessary to draw some conclusions concerning theoretical sociology at least in this systematic final chapter of this study, and to point to the concrete research program that follows from the considerations and results presented here. As a matter of course, there are a lot of prospects, links to other fields of research, and especially open questions that arise from what we have said so far. To mention them all would require an extra study. So, for mere shortage of space, I have to confine myself to picking out a few aspects that are particularly important. In sequence, we are going to deal with the classical problem of the social rule; the not-so-classical problem of social universals; and a listing of the most important open questions that a mathematical sociology built up on these fundaments will have to (and is able to) deal with. So, strictly speaking, the actual sociological work only just begins with the ending of this book. Yet, the results we have attained so far already show that it really makes sense to expand that above-mentioned famous remark of Kant from the natural sciences into the social sciences.

Let us start with a summary of the fundamental principles, procedures, and initial results:

[107] Vorrede zu Metaphysische Anfangsgründe der Naturwissenschaft. Immanuel Kants Werke Band IV. Berlin 1922, S. 372: Cassirer. (I postulate that in each particular philosophy about nature is only as much science as there is mathematics.)

J. Klüver, *The Dynamics and Evolution of Social Systems*
© Springer Science+Business Media Dordrecht 2000

(a) The basic sociological concept chosen for the procedure in question is that of the *social rule*. This means, in particular, that social systems are characterized as rule-constituted systems of actions and interactions, and that their self-organized adaptive dynamics has to be understood as a result of the impact and/or variation of their relevant rule systems. Thus, our theoretical starting point is not a single social actor, but the rule-governed interaction between at least two actors. This leads, on the one hand, to a resolution of the classical difference between systems theory and action theory and between macro- and micro-theories in general; on the other hand, it allows a deliberate theoretical junction to theoretical positions that have traditionally criticized the individualistic bias of starting out from the single subject - from Mead and symbolic interactionism to Habermas. We are going to deal with this in more detail in section 5.1. The arguments against a methodical/theoretical individualism are sufficiently known, and we do not need to repeat them; instead, I may confidently point to the "intersubjective" tradition here.[108]

Nevertheless, this option in favor of the rule concept as the theoretical basis is not a matter of course, neither in the natural nor in the social sciences. That is why this option will be substantiated and explained in more detail in section 5.1.

(b) The *general* basic theoretical concept that enables sociology to achieve its discipline-specific way of joining to the new sciences of complex systems is that of the *adaptive complex system*. This is to say that social systems always have to be understood and modeled as systems that vary their interaction rules by meta-rules so that they can adapt themselves to environmental requirements through self-organization. Note that the relation of meta-rules to base rules - or of meta-meta-rules to meta-rules and base rules - is not an unilateral one, but that there is a mutual interdependence: the meta-rules can vary the base rules, and can likewise be varied and even be annulled by the dynamics of the base rules. In the very illustrative words of Hofstadter: logically, adaptive systems can be characterized by strange loops that are effective through the single levels of rules, meta-rules, and so forth.

In these layerings of different rule systems we can recognize the central *logical* difference between the systems of the natural sciences and those of the cultural and cognitive sciences. "Natural" systems usually need only one rule level, and in this respect - though only in this respect - they are "simple" in comparison with social systems. Therefore, the

[108] From a sociological point of view it is thus somewhat disturbing that the individualistic actor's perspective is quite prominent especially among recent approaches to computer modelings of social processes. This might easily lead to pouring rather old wine into new wineskins, so to speak, which would be a theoretical step backwards for sociology as a whole.

sociological use of concepts from natural scientific contexts, like that of adaptive systems, must not be confused with natural scientific reductionism: the peculiarity of the social is based on this difference and must always be kept in mind, especially when trying to conceive a theory of social systems with strict concepts and methods.

Though the introduction of different rule levels might seem quite complicated at first, it obviously makes it possible to comprehend genuinely social - and "mental" - phenomena such as learning, memory, mind, or, in a word, any kind of rule variation with the desired precision and to examine them by means of formal systems and models. We could also see that this even allows exact analyses of the concept of social space and the evolutionary unfolding of dimensions.

(c) *Methodically*, this is achieved by the construction of formal systems, i.e., computer programs that are hybrid in the technical sense of the term. We only have to take a look at social systems to see that meta-rules usually have a different form than the corresponding base rules; therefore, it suggests itself for logical as well as economic reasons to model adaptive systems with programs that are made up of various different programs.

The construction of such programs is done in the following way, according to the principle of "cybernetic explanation" formulated by Bateson: first, a general model is developed that has much greater degrees of freedom than the "real" system that is to be modeled; subsequently, the empirically known constraints are introduced in order to make the model's behavior as empirically adequate as possible. The point here is not to model one single system - at least not in this study - but instead to model a specific system class. For example, the purpose of simulating the evolution of different levels of differentiation was not to model a specific society, but to model the process of differentiation "as such". However, the results we obtained from this rather general procedure were already quite similar to real processes; this supports the assumption that these processes follow a general logic that can be described with the concept of "social universals" (see below, section 5.2.).

(d) Both the dynamics of logically "one-dimensional" systems, that is, systems with only one rule level, and the dynamics of logically hybrid systems are obviously subject to particular regularities which can be characterized with the concepts of the classes of complexity, the adaptive success, and, first of all, the control and meta-parameters. Logically one-dimensional as well as logically multi-dimensional systems generate system dynamics the properties of which can be put down to the rule systems' characteristics determined by specific parameters. Apparently,

this is true for both self-organizing and adaptive systems, though in the latter case mainly the described meta-parameters are important while the classes of complexity do not play such a decisive part. Though our results are still rather preliminary we can discern some mathematical regularities in the apparent jumble of social changes.

The terminology of rule-based adaptive systems seems to make it possible to search totally different phenomena for similar regularities, namely for the values of meta-parameters. It is just this terminology that also makes it possible to show similarities between memory, mind, and comparable system properties: they can be characterized by the variation of particular meta-parameters, and thus as system mechanisms that make such variations possible. The results presented in section 3.3.3. concerning the connection between adaptive success and the values or variability of meta-parameters demonstrate that this terminology really allows a reasonable procedure, and that important insights can be obtained with regard to the behavior of adaptive, i.e., hybrid systems.

These are of course results about "pure" systems, and they do not say anything yet about concrete social processes (see below, 5.3.). However, according to Bateson's conception of the cybernetic explanation we are able to derive something about "real" systems and their concrete behavior from our results, namely that they are subject to the mathematical constraints of logically multi-layered rule systems and that their actually performed behavior is *the only mathematically possible behavior*, taking into account particular biological, ecological, and other constraints. In a way, real systems behave according to the famous motto of Sherlock Holmes: if all other possibilities are eliminated - that which the constraints do not permit - then what remains has to be the truth, in other words: the explanation for the real behavior. This also goes for our geometric considerations of section 4.4.: a system does not necessarily get beyond the level of segmentary differentiation; as we could see, it is even rather improbable that this evolutionary level should be achieved. However, once it is achieved, e.g., because of particular ecological and demographic constraints, the system has practically no choice but to behave in the way that was analyzed with TRISOC - including all the resulting problems. Thus, the methodic procedure presented here can be justified by the frequently mentioned Kantian stance: the mathematics of pure rule systems characterizes what has to be regarded as the condition of the possibility of social complexity, and also explains it.

What this little summary shows is that it is indeed possible to make some progress on the path presented here, and to hold out quite concrete prospects for the conception of a sociology that is both ambitious as regards content and methodically strict, that is, proceeding mathematically. Nevertheless, I have emphasized at the beginning of this

study that I am not presenting a theory, but building blocks and guideposts. The next steps on the way towards such a theory ensue not least from the problems that could not (sufficiently) be dealt with so far. Examples of the most important ones shall be presented in the following subchapter.

5.1. THE REGULARITIES OF SOCIAL ACTION: SOME ASPECTS OF THE SOCIAL RULE CONCEPT

> "For what a thing means is what habit it involves" (C. S. Peirce, The pragmatic maxim).

At the beginning of this study, I have explicitly pointed to the fact that the *rule* concept is the *logical* center of the considerations and results presented here. At the same time, it was presupposed that the sociological tradition can be adequately described and reconstructed with that concept. I would like to emphasize once again that this premise can neither be taken for granted in computer modelings nor in systems theory approaches: both research tendencies, which are converging more and more, originate in the tradition of the natural sciences, which has been determined since Kepler and Galileo by theoretical modelings in the form of equations as general as possible. Despite the earlier mentioned game theoretical analyses in evolutionary biology by Maynard Smith, most ecological and evolutionary systems theorists still work with modelings of biological systems by means of differential equations according to the classical Lotka-Volterra equations (cf. Pimm 1991) as far as I know. As a matter of fact, modelings by means of equations are also used in simulations of economics and social sciences (e.g. Troitzsch 1995, Leinfellner 1995, Epstein 1997).

The reasons for preferring the concept of the rule in this study have already been mentioned in chapter 4.; they can be systematically summarized like this:

(a) In the vast majority of cases it is just not possible to describe social connections and processes with equation systems. The history of mathematical sociology has always shown that only rather simple processes, which are usually not very substantial for more ambitious sociological questions, can be formally grasped in the classical mathematical way (see, e.g., Beltrami 1993). However, it is possible to reconstruct the *logics* of social processes, and in the next subchapter something will be said about that; on the other hand, it is usually impossible for empirical and methodical reasons to describe such reconstructions formally as differential or difference equations. The

undeniable successes of economics, where exactly these modeling techniques are used, only show that this just works on condition of the existence of indisputably quantifiable units. Even the success of economics is known to be limited, and this may well be due to the narrowness of its formal methods.[109]

(b) Empirically, by contrast, it is in fact possible to grasp social processes by reconstructing the rules according to which social actors behave or have behaved. This is *the* classic subject of qualitative social research, and its fruitfulness has been successfully demonstrated again and again from Mead and Weber, through Schütz, to Berger and Luckmann, as well as Garfinkel. *Empirically*, the use of the system concept in the social sciences only makes sense as regards action systems, and thus systems whose basic logical concept is necessarily that of a rule. The originality of social systems, as against the systems of the natural sciences, is expressed here both theoretically and methodically. It is true that neither Parsons nor Habermas, who feel clearly obliged to the action system concept, have explicitly picked out the rule concept as a central theme; however, in the case of such profound experts on Weber, we may justifiably assume that they simply presuppose the rule concept, and do not think it necessary to make an issue of it. This is evidently not without its problems; anyway, the only possible way of joining the considerations presented here adequately to the classical tradition of both empirical and interpretative sociology is by focusing on the rule concept.

(c) The main reason for preferring the rule concept, after all, is the repeatedly mentioned fact that social systems cannot simply be characterized by "one-dimensional" dynamics on the basis of equally one-dimensional rule systems (see section 3.3.); they instead have to be described by their *variability*, and that means adaptivity. Grasping this basic feature in a formal model will only work if the form of variable rules is employed. Of course, one can also change the parameters of equations that have been formulated for the description of the system behavior; likewise, it is technically possible to maintain the logical structure of equations variable. But such changes would be so forcible that it would be very difficult to see any meaning as regards content in them (see also Read 1990).

At this point we can also see the limits of pure time series analyses, which are often used to obtain equations for the description of system behavior in the quantitative social sciences (e.g. Eve et al. 1997). If the

[109] The general presupposition of a rational *homo oeconomicus*, which usually seems rather strange to non-economists, can only be pointed at as an additional problem here. Waldrop (loc. cit.) gives a quite illustrative description of the disbelief of the participating physicists at the first conference of the Santa Fé Institute when they asked their economic colleagues: "And you really believe that?"

history of a social system is reconstructed only from the changes of particular statistical values, and if this is taken as the basis for deriving general regularities, e.g., in the form of equations, then one can never rule out that specific rule variations might have happened within the system during the observation period, which subsequently generated the observed values. Hence, time series as such, like any statistic, are not very meaningful per se; only a historical analysis of the system with regard to its rule variations, and thus with regard to its actual variability, can give information about how the system behavior - as regards the analyzed time series - is to be explained. This explanation, to stress the point once again, can only consist in measuring the variability of the system - the formal part of the explanation -, and in postulating the reasons for this variability - the empirical part. Only such an explanation may allow a statement - with all due caution - on the probability of the system's possible further behavior.

Since the appearance of chaos theory and the broad interest in non linear systems, which was partly caused by that, it has been repeatedly stressed that linear extrapolations of an observed system behavior are impossible - in whatever form this may be done. Our considerations about the preference of the rule concept show furthermore that it is probably rather futile too to look for non linear extrapolations in the form of appropriate equations in the case of social systems. Another approach seems to be more promising and capable of being joined to the more interpretative sociological tendencies: the study of the history of a social system in the classical sense of the word, that is to say, the history of its interaction rules and its rule variations, in order to derive from that its variability and determine its probable future development. If a system has shown only low degrees of variability in the past, and if it seems empirically reasonable that conservatism is part of the system's social identity, then it is possible to state the system's future variability, and thus its future history, with a specific probability. Historical analysis and formal system measuring apparently enter into a necessary symbiosis here, although the prognostic statements gained from this) can of course only be formulated with great caution. Anyway, such a procedure would neither be a futurological way of reading tea leaves, nor a reduction of social complexity in Procrustean beds of the natural sciences.

This is not to hypostatize a logical contradiction between rules and equations, however, which does in fact not exist. Rules can usually be written as equations in computer programs; many of the logical rules of our CAs, e.g., are formulated as matrix transformations and thus as algebraic equations. It is just a question of which technique actually makes sense. Likewise, equations can be interpreted as rules, which is

how they are often formulated in algorithms: on the one hand, the delta learning rule in 2.3 is formulated as an equation, on the other hand, it "is" nothing but a rule for the program to change the weights of the respective neural net systematically in the specified way. The rule concept we are dealing with here is preferred for methodical reasons and it has the advantage that it can be joined to the traditions of sociology - no more, no less. As always, the fruitfulness of this approach must be measured by its successes; new ways to measure cultural changes by Eigen-Schuster equations seem to be another possibility (e.g. Leinfellner 1995).

Following these methodical remarks, we will have to argue in a somewhat more sociological way as it were, and question the presupposition according to which sociology can be characterized as the reconstruction of social rules. After all, the concept is not *that* central, and we cannot simply take this statement for granted. Sociologists speak of structures - whatever that may actually mean -, of institutions, of social nets, or also of systems; of course, this is not even a beginning of a list of the central sociological concepts. Compared to this, the rule concept is not nearly as prominent as far as I can see, and it is probably no accident that, e.g., there is no systematic explication of the concept of the social rule in Luhmann.[110] So let us take a closer look at that concept.

There are presumably two significant reasons why the rule concept is systematically neglected in theoretical sociology: on the one hand, it could hardly be seen until now which reasonable questions of research (and results) might be gained from that concept; the systematic analysis of (social) rule systems is certainly only fruitful in connection with the methods described in this study. Only the mapping of social systems and their rules into formal systems of the kind described here opens the possibility of systematically analyzing different social rule systems with regard to their formal properties, and deriving therefrom both the real and the possible behavior of social systems. On the other hand, the concept of the social rule is in fact quite ambiguous; so far, we have presupposed the premise that it is always unambiguous to speak of social interaction rules; this premise is at least not trivial, and has to be continually proven in concrete cases. This led, e.g., Bourdieu to characterize the rule concept as unsuitable for sociology: "I think that the language of the rule is often a hiding-place for ignorance in the social sciences." (1989/1991, 94) This statement is all the more surprising because its author, when he conceived his habitus concept, was guided to a great extent by Chomsky's conception of a grammar as a formal rule system.

[110] The well-known definitions by Giddens (1984) to characterize institutions as sets of rules are still an exception.

Bourdieu comes to this opinion as a consequence of his critique of Lévi-Strauss' conception of structures and rules, which he regards as too rigid; according to Bourdieu, Lévi-Strauss suggests that social rules like, e.g., marriage rules work analogously to laws of nature in the sense that social actors always and almost compulsively orient toward them. In real life it is much more likely that actors behave "strategically" (Bourdieu) in particular situations, namely by organizing their actual behavior variably, taking situational peculiarities into account, ignoring or breaking existing norms if necessary, etc. For Bourdieu there is no such thing as *the* rule, but one has to distinguish at least between rules that are official norms and *regularities* of actual behavior. The latter do not generally coincide with the obeying of official norms, but ensue from strategic acting that depends on the situation.[111]

The realization of the ambiguity of the social rule concept is of course not new, and even Max Weber (1982/1907) has already worked out this point in great detail. He uses the example of the card game "skat" (comparable to piquet) to illustrate the difference between *political* rules (of skat) - which card combinations are the most valuable -, *legal* rules (of skat) - whether a player has lost a game if he violates a particular rule -, and "*Erfahrungsregeln*" (empirical rules) - when it is a proper strategy to underplay in order to catch the 10 (loc. cit., 337). Like Bourdieu, Weber stresses that it does not make much sense to speak of social rules without further differentiation; the real acting of social actors with regard to rules can of course not be understood without stating the type of rules they belong to. Taking part in an interaction system "game of skat" requires the absolute obeying of the political and legal rules - which may have been explicitly codified before the game - otherwise one is not playing skat at all. In this respect, these rules constitute and define the social system "game of skat"; accordingly, they are the "*Voraussetzungen*" (conditions) (Weber loc. cit., 340) of the possibility to recognize an interaction system as a "game of skat". A violation of these rules either leads to an end of the whole interaction system, or it is sanctioned like any other violation of legal or political rules - for example, by loss of the game with 120 minus points.

Not obeying the empirical rules, on the other hand, may cause disadvantages in a single concrete game - in particular, the opponents will mock the blundering actor and the other player will blame him respectively - but it is not regarded as an official rule breaking. One might even say that an interaction system such as skat is virtually

[111] Bourdieu's critique of a too simple rule concept largely parallels the critique of rule-based AI research by Dreyfus and Dreyfus 1987. I thank J. Reichertz for drawing my attention to the relevance of this text of Bourdieu.

calculated for the fact that the single players follow such empirical rules with a different frequency, and particularly with a different adequacy to the situation. It is this aspect of individual skill that is known to make the difference between skat and games of pure chance; the various skat columns demonstrate that skat can be analyzed and handled in a game theoretical way, just like chess.

We can generalize this example, according to Weber, and what ensues is again Bourdieu's differentiation between official norms and practically obeyed rules, with the latter being valid and obeyed only with regard to the particular situation. At the same time, this shows that at least these two types of rules must exist: on the one hand, there are general norms, which - according to Weber - have a legal-political character and are constitutive for a social system; they determine what is actually the matter, and they make it possible for an observer or an actor to recognize that. This does not mean that the individual actors necessarily behave according to these norms in every single situation or, in reverse, that their acting could be derived and thus explained from these norms in every situation. Additionally, there must be situational empirical rules (Weber) or strategies (Bourdieu) that add concrete content to that general frame as it were. This coincides with what we said about the connection between the theory of games and systems theory: general frame rules are necessarily "underdetermined" in the sense that they only indicate what is not allowed and which options remain in a concrete situation of decision. The *concrete*, particular decision, however, does not follow therefrom, but has to be made by the use of additional aspects, including individual experience and a "feeling for the game" (Bourdieu) as well as game theoretical analyses.[112]

The differentiation between general frame rules and single situational rules is of course much too rough to comprehend the complexity of "regular" social reality. Frame rules need to be carefully distinguished themselves with respect to their relevance for the system and their actual binding force, to name only one aspect. For example, the Catholic Church's general ban on contraceptives undoubtedly belongs to the frame rules that constitute and thus define the social system "Catholic Church". Nevertheless, its actual binding force is probably not very strong, at least not in European countries; in a formal modeling of this social system, the rule "ban on contraception" could only be calculated with a rather low

[112] Bourdieu's rejection of "juridism" (loc. cit., 86) is an inappropriate reduction itself: it is true that actors in particular situations do not necessarily obey the general norms and that the latter are not sufficient for an explanation; but one can not conclude from this realization that those norms are irrelevant - neither for the actors nor for the sociological analysis. The sensibleness of concrete situational strategies is in fact determined by the general frame as well. One always needs these two types, both practically and theoretically.

probability parameter. In this case, a general rule is permanently counteracted by individual action strategies. Compared with that, a frame rule like, e.g., that of the confession presumably has a much stronger binding force. (I have to leave it up to theologians whether or not the ban on contraception is a fundamental rule within the social system Catholic Church.) Anyway, this simple division is perfectly sufficient for the purpose of practical research, and especially for the analysis of social systems by means of formal modelings.

If social change is defined primarily as rule variation and rule extension, as is done in this study, then several possible mechanisms of rule changes ensue from this conceptual distinction: the first possibility is mainly exercised by the hybrid systems presented here, namely the rule variation for an entire system on the basis of system adaptation. By that, the system is regarded as a kind of collective actor who constitutes himself in a totally new way with new overall rules by generally binding political decisions. A second possibility results from the fact that individual actors virtually annul single frame rules in their practical acting. In the Federal Republic of Germany this happened, e.g., in the field of sexual behavior; the system as a whole could only modify the frame rules subsequently in order to make them fit the actual particular rules. A third possibility arises from the fact that the frame rules, rather than being directly annulled, are "stretched" to such an extent that the single actor's action space is in fact hardly restricted. This is usually only the case with religious or political "formulas of orientation" where certain basic values are regularly referred to, but only situational rules are actually followed.

The modelings presented so far only concerned "system rules" without further differentiations, and so it was not necessary to take the above-mentioned distinctions into account yet; this should be made up in future studies (see below, section 5.3.). However, our game theoretical considerations could already give a suggestion of how this might be achieved without meeting with fundamental problems. If, and to what extent, that is required depends of course on the actual question; the fact that frame rules are neither sufficient by themselves nor unrestrictedly obeyed has already been allowed for by letting many of the presented models operate with stochastic system rules. From the actor's point of view this simply means that the single actors have individual action spaces.

If we think of the terminology of adaptive hybrid systems that was developed in section 3.3, the distinctions of Bourdieu and Weber can be taken up quite easily: their meaning for a system is obviously that some of the totality of rules that are relevant for the system have to be kept

constant, while others can be varied according to the situation. In other words: the social actors that Weber and Bourdieu had in mind have to be taken as adaptive action systems themselves, and their r parameter is not equal to zero, but also is not equal to one. The strategic players in the GA model of the prisoner's dilemma are obviously able to behave in a highly variable way to find out their respective optimal strategy, but on the other hand, they have to keep the pay-off matrix constant. The players in the model of Nowak and May are tied to the locality of their neighborhood, and while their behavior does indeed depend on the situation, they always behave in the same way in particular situations. So here we have a low variability and a clear predominance of the rules that transcend the situation. Finally, the variability of social mobility parameters in systems like SOZION conveys, on the one hand, the actual rules' dependence on the prevailing environmental requirements - the situation; on the other hand, a constancy of more general rules is characterized by the constant and steady operation of the genetic operators on those parameters. Apparently, the formal systems described here always ran, strictly speaking, in accordance with the distinction that was stressed by Weber and Bourdieu, and they were doing so because of purely formal considerations; however, what Weber and Bourdieu did not (yet) take into account, in contrast to our considerations, is that social evolution can also mean that *all* rules can become disposable in principle - in the case of $r = 1$.

So, if sociological analysis should focus on the concept of the rule and, in consequence, sociological knowledge is understood as a reconstruction of social rules, then the question arises how "complete" the outcome can possibly be. Structurally orientated traditions of thinking would probably not have any difficulties in generally accepting such an understanding of the tasks of sociological research, despite the mentioned differences, e.g. between Lévi-Strauss and Bourdieu. Scientists with an ethnomethodological inclination should also feel on familiar ground; Weingarten and Sack express that very precisely: "Das Anliegen der Ethnomethodologie ist ein ganz ähnliches [wie das der Ethnologie, J.K.]: Ihr geht es darum, die Methoden aufzudecken, deren sich die Gesellschaftsmitglieder bedienen, um die Vielzahl ihrer Alltagshandlungen durchzuführen" (Weingarten/Sack 1976, 10).[113] If we just replace "methods" with "rules" this resembles exactly the point of view that is supported here.

It has to be said, of course, that most followers of the many varieties of interpretative social research, including ethnomethodologists, would not consider methods as identical with rules of the kind discussed here,

[113] Ethnomethodology tries to discover the methods which social actors use to accomplish their everyday actions.

but the actors' *knowledge* of the rules plays a fundamental part too. This aspect has not been dealt with here. On the other hand, however, it was shown again and again by the works of Oevermann (e.g. 1991) and others that social actors can follow rules even if they are not aware of them, and even if it is impossible to become aware of them at all. It is well known that this fact has led Apel and Habermas to postulate the necessity of a form of social sciences that reconstructs in a quasi-natural scientific way the social rules and regularities which the actors cannot recognize immediately, that is, from the action perspective of their everyday life (e.g. Habermas 1981 II; Apel 1979).

Thus, empirical research has to distinguish between the rules that the actors are aware of (and possibly need to be aware of) and those that they either do not have to be or can not be aware of. In the case of the interaction system "skat", the actors probably have to be conscious of the essential rules, or at least they can become conscious of them in situations of decision. A proficient player would of course not consciously think about underplaying in order to catch the ten in a particular situation of the game; he does that automatically, as it were, on the basis of his long playing practice and the empirical rules he has acquired thereby (cf. Dreyfus and Dreyfus loc. cit.). Yet, he is able at any time to explain his behavior on demand with exactly these empirical rules. It is in fact the methodical premise of every ethnologist and interpretative social researcher *that* social actors are usually able to give such explanations. In contrast, no one would expect a lower class adolescent to give an explicit explanation of the rules that brought about his/her particular occupational biography. He or she would probably declare that he/she just did this and that "because that's just the way things are done."

So the question of whether or not the actors are aware of particular rules is secondary as regards aspects of modeling and the related epistemological interest, however important the question is with regard to empirical procedures. Anyway, there is no reason to assume a difference between what can be comprehended by reconstructing the rules of social acting and what can be generally said about the respective social processes *from a sociological point of view*. This point can not be set out in great detail here; however, if it is true, then formal models of the kind presented here are also an excellent instrument for validating hypotheses about reconstructed rules: the latter are mapped into a formal system, and then the system's dynamics is compared to one's actual observations. Hence, we can not only affirm that computer modelings and interpretative empirical social research do not exclude each other, but that they obviously complement and need one another in a

fundamental way: the first become empirical by the second, and the second becomes as precise as one can possibly wish for a discipline that is often rather vague. Of course, this mutual complementarity holds *a fortiori* for an empirical social research that is rather quantitatively orientated.

The central position of the rule concept that is postulated here for both sociology and computer modelings renders some additional clarifications as well - as a bonus, so to speak:

(a) In the terminology of systems theory, which was used primarily throughout this study, the iterated application of interaction rules to an initial state is known to yield the dynamics of the system, that is, its trajectory or sequence of states. Variations of these rules and/or the initial states produce different trajectories and so forth. If we interpret these rules as action rules, then their continual application, that is to say, the actor's compliance with them in a situation of action, has specific consequences - regardless of whether these are intended or not (Knorr-Cetina 1981). Thus, in terms of action theory the dynamics of a system are just the chain(s) of intended or unintended consequences of the following of particular social rules. It becomes immediately obvious that, and to what extent, action theory and systems theory turn out as two different descriptions or wordings of the same object - the dynamics of social processes - by means of the rule concept. Therefore, it seems about time to abandon this venerable dichotomy once and for all.

The *Unübersichtlichkeit* (difficulty in understanding) of complex action contexts and their consequences in modern societies that was stressed, e.g., by Habermas (1981 II) is nothing to be astonished by if we examine it from this aspect: it is simply an expression of the systems theoretical realization that particular rule systems can have trajectories which cannot be grasped by intuition or by means of ordinary patterns of thought (see, e.g., Dörner 1992). However, this is a characteristic feature of modern societies only in that the latter most likely have rule systems with values for the respective control parameters - e.g., in the case of the λ parameter, relatively easy access to all social roles - which generate rather complex dynamics; and, of course, in that they show a high variability of the rules, which corresponds to high values of the meta-parameters. This should be intuitively clear; we have to leave it up to future research to produce more detailed studies on that. As a result, modern societies are only "*unübersichtlich*" if one tries to comprehend them by informal means; from the point of view of formal mathematical rule analyses, the matter is not unclear at all.

(b) The no less venerable opposition between structure and process, which has already been rejected by Luhmann and others, becomes invalid as well. Traditional structuralist approaches are usually based on

a too narrow concept of structure. If, for example, institutions are interpreted as classic forms of social structures, a point that should be largely agreed on, one can easily see that institutions are in fact also nothing other than rule systems; the notorious conservatism of institutions points to the fact that these rule systems can be characterized by trajectories that usually have an attractor of the period zero. Hence, social structures, inasmuch as they express something static, are nothing other than rule systems with simple attractors - with the period zero in the extreme - and with very low meta-parameter values. Again, we obtain static structures as the mathematically simple borderline case of complex dynamics, and that's all. The (pragmatic) "meaning" of social structures, to adopt the initial quotation by Peirce, lies in their action rules, especially the geometrical ones, which define and constitute these structures, and thus, in the "use" of specific rules by social actors. Pure structural analyses like the analysis of social networks undoubtedly make sense; however, one must not forget that what are analyzed there are borderline cases of more general phenomena, and that there is no contradiction between structural analyses and analyses of processes.[114]

(c) There have always been attempts by natural, as well as social, scientists to transfer concepts from (biological) evolutionary theory to socio-cultural evolution - which is in fact where that concept comes from. The trouble with that was always the question of what should be the corresponding social items of the genes and the genome, respectively, that function as the vehicles of evolution through variation, selection, and reproduction. The biologist Dawkins, e.g., postulated the existence of so-called "memes" in this context (Dawkins 1976), without, however, giving a precise sociological explanation of this beautiful concept.[115] The theoretical sociologist Giesen (1980) brought in Popper's "world 3" to solve the problem and put "ideas", "thoughts", etc. in the socio-cultural place of the genes.

By introducing the social rule as the basic evolutionary unit, which we regularly did in the computer modelings described in this study, those evolutionary approaches in sociology lose their vagueness all at once, and can obviously be dealt with in a precise way. If one bears in mind that socio-cultural evolution proceeds according to the model of Lamarck

[114] Of course, such a pragmatic or operative definition also concerns religions, conceptions of the world, systems of thought and faith, and so forth. Their meaning, inasmuch as it is a *sociological* one, lies in the rules that are derived from those views of life and in the behavior that is induced by these rules. So, from a sociological point of view, such edifices of ideas *are* nothing other than specific ensembles of rules.

[115] Meanwhile there is even a journal that is exclusively dedicated to the investigation of such memes, which can be found in the internet.

rather than according to that of Darwin, in that in socio-cultural evolution, in contrast to the evolution of nature, it is possible to "inherit" learned and otherwise acquired characteristics, then social evolution shows itself perfectly parallel to biological evolution as the variation, recombination, mutation, selection, and reproduction of rules and rule systems, respectively. We do not need to go through the corresponding examples from chapters 3 and 4 again. So, we neither need excursions into a "world 3" of objective ideas, nor hypostatizations of memes and other cryptic units; what we need is just the good old concept of the social rule. With that, an immediate orientation to models of the theory of evolution obviously becomes possible and - for the sake of clarifying notoriously vague concepts such as "social change", "historical processes", etc. - also necessary. Again we see, though, that sociology has to develop its theoretical concepts for itself, and that it cannot simply rely on literal borrowings from the natural sciences. Otherwise, one quickly ends up with dubious concepts like sociobiology, which has meanwhile disappeared from the discussion - and rightly so.

Another traditional opposition is that of internal and external factors of socio-cultural evolution. This opposition is also not necessary if one defines adaptability as the competence of rule variations. Of course, rule changings always occur as a result of environmental demands - the external part of the evolution of the system. Yet, the changings and extensions of the rules are also always achievements of the system, and therefore internal factors - if one needs this concept at all. The environment determines that, how much, and in which directions, adaptive processes are to be realized; how the system reaches successful variations is up to it alone.

Now, if sociology is the systematic reconstruction of social rule systems for the aforementioned reasons, then the possibilities of a strict, i.e., mathematical sociology are evident, at least as a program: inasmuch as sociology is empirical, it is concerned with the reconstruction of the rules according to which social actors behave in particular contexts, and which constitute these contexts; it is also concerned with reconstructing the ways in which these rules are varied. The results of these reconstructions can then be formally modeled and analyzed as a formal rule system - with regard to their parameters and the trajectories and rule variations generated by them. This may lead to more general regularities, which must not be misinterpreted as quasi-physical laws of a "second nature" (Lukacs, see below), though. In any case, it will lead to a mathematical description and analysis of social processes, which would offer the possibility of telling what is empirically contingent about social processes, and what follows from the formal logic of complex system dynamics. That would already be something.

So, *methodically*, there is every reason to start out from the social rule as the basic concept; as regards the *object* of sociological research, the central importance of the rule concept is valid as well for any approach that does not start out in a "monadological" way from the solitary individual, which never exists in reality. This is expressed by Wittgenstein's famous dictum according to which "*einer allein nicht einer Regel folgen kann*" (no one can follow a rule alone) (Philosophische Untersuchungen § 202); the existence of a rule immediately points to the social, namely to a social context in which, and for which alone, rule-governed actions are the archetype of sociality. On the other hand, Berger and Luckmann (1966) show in their frequently mentioned classic that societies come into being by the introduction of rules, and that they develop by the stabilization, transference, or even changing of these rules. Thus, by following rules, people "construct" (Berger/Luckmann) their society; on the other hand, society is in practice always presupposed by introducing rules. Following both Wittgenstein and Berger/Luckmann, we can certainly say that the social rule is the practical condition of the possibility of any sociality, which leads us into the context of the "transcendental pragmatic" reflections of Apel and Habermas. The concentration on the rule concept, which was instigated for methodical reasons here, can obviously be legitimized epistemologically in any respect that might be desired. But this leads us on to the next subject.

5.2. SOCIAL UNIVERSALS AND (BIOLOGICAL) CONSTRAINTS

If one follows our proposal to analyze social systems by mapping them into formal systems and studying their behavior by experiment, then it obviously suggests itself to ask for *general* properties of social systems, that is to say, common basic patterns underneath the diversity of social reality and realities. This has already been mentioned in chapter 1.; let us now look at the inherent possibilities more systematically under the heading of *social universals*.

As far as I see, the concept of universals has been introduced into the recent scientific discourse by Chomsky and his postulating of "linguistic universals" (Chomsky 1965). The term 'universals' is of course much older, we only have to think of the Scholastic controversy concerning universals; as regards the concept of "social universals" and its use in the history of science, there is an informative survey by Brown (1991). It was Chomsky, however, who gave this concept a new topicality, by which he established an almost paradigmatic orientation, and not only in

linguistics. He distinguished between two kinds of linguistic universals, namely *formal* and *substantial* universals:

> "Substantial universals such as these [e.g. semantical categories or phonetical units, J.K.] concern the vocabulary for the description of language; formal universals involve rather the character of the rules that appear in grammars and the ways in which they can be interconnected" (1965/1969, 29).

Substantial universals, which are often taken as linguistic universals in general, can be found, e.g., in the fact that all languages have different categories of naming persons, or that a distinction is made between noun, object, and verb of a sentence, which are the syntactical basic units of sentences. These could be called "absolute" universals, because they form the framework of any language. Besides that, there are "implicative" universals, which were already mentioned above: a famous example for this concerns the fact that one distinguishing feature of languages is that they differ in their word order of noun, verb, and object (NVO). German and English are NVO languages, while Japanese is an NOV language. This fact implies the following implicative relation: if a language has an NVO word order, then interrogative clauses are formed with the interrogative particles standing at the beginning of the sentence, and the language has prepositions; if, on the other hand, it is an NOV language, then the interrogatives are at the end of the sentence and it has postpositions. Pinker (loc. cit.) mentions in this context that most of the substantial universals have an implicative nature, and that hundreds of such universals have already been found.

In comparison with that, formal universals cannot be described as easily, and they explicitly refer to grammar as a logical ensemble of formal rules. The examples Chomsky (ibid.) gives in this context concern the fact that grammars contain transformational rules for the mapping of (semantic) deep structures into (phonetic) surface structures; another example is the recursive nature of transformational and generative rules (cf. e.g. Grewendorf et al. 1987). In his classical work, from which the above quotation is taken, Chomsky does not explain this any further; however, the reader of this study should be familiar with a definition that has since been used to refer to the properties of rule systems. We shall return to this point later.[116]

Chomsky's postulate of linguistic universals, and the corresponding conception of a universal grammar, also attracted attention in the social

[116] I was pointed to these passages in Chomsky's work by Helga Andresen, who also drew my attention to the fact that the theory of generative transformational grammar has of course been developed further in the meantime both by Chomsky himself and by his school (see also Pinker loc. cit.). Being a non-linguist myself, I can only refer to that here.

sciences. The (cultural) anthropologist Brown (1991), orientated to the concept of substantial universals, made an attempt to make Chomsky's terminology fruitful for the social sciences. Brown analyzed and critically re-analyzed - see the above remark on the mistakes of Margaret Mead - all available ethnographic descriptions of different societies; from that he distilled general features which could be found as basic patterns in all, or at least most, of the analyzed societies. By that he obtained very extensive lists of features that are common to all societies; in addition, he lists a lot of features which can also be characterized as universals with a significant statistical frequency, e.g., the predominance of men in the public sphere. This corresponds to the linguistic use of the concept of (substantial) universals, because statistical frequencies are taken into consideration there as well. Let us pick out some illustrative examples:

In every society there are laws that make murder, rape, and violence in general a punishable offence; all societies have institutions for the mediation of conflicts; the institution of property is also universal; the same goes for "institutions of leadership", which are not organized dictatorially in most cases; sexual interactions are usually - as a social rule - closed to the public; the same is true for excretions and their removal; there is almost always a differentiation of the society members according to status and prestige; the institution of heritage is another universal feature, and so forth (Brown loc. cit., 130).

Brown's result seem to stand in contradiction to a number of favorite myths of the social sciences: there are neither indications of primordial matriarchies, nor of early communist societies; moreover, people seem to have always dealt with nature in the way that is discernible in Habermas' remark (1968) on work as an "anthropological constant": nature as available matter. The contrast between society and nature is in fact another feature on Brown's list of universals. The main point, however, is that the common view of man being a creature that can be varied by society in almost any way - Gehlen's *Mängelwesen* (deficient being) - becomes rather doubtful by Brown's conception of "Universal People": apparently, people have organized their social systems at all times, and under very different circumstances, in a very similar way.

Pinker (loc. cit.) rightly warns of taking Brown's results as a list of instincts or innate psychological inclinations. The history of modern societies has particularly shown that although not everything is variable about social rules, new rules are constantly put to the test, some of which pay off, while others do not. However, the undeniable variability of social systems, which immediately reduces any socio-biological reductionism to absurdity, obviously does not bring all logically possible

and thinkable societies into being, but instead regularly produces results which can indeed be described as substantial social universals - a fact that, in addition to Brown's formulation of the Universal People, has to led to the concept of a "meta culture" (Sperber 1982), that is, the entire common conditions of all different societies. Evidently, the substantial social universals are no less empirically proven than their linguistic counterparts. Yet, in order to avoid a biologistically simplifying explanation of the former according to the ominous motto "one gene for every social universal", one will have to ask for formal universals: thus, we have to deal with formal characteristics of social rule systems, which may possibly explain why particular substantial social universals necessarily had to arise, and why social systems impose constraints on themselves in this respect, that is, why they did not realize other possibilities.

The analyses of formal systems in chapters 3 and 4 have already indicated what the search for the formal social universals might look like, and what character such universals might possibly have: we learned that the dynamics of "logically one-dimensional" systems, i.e., systems with constant rules, depends on the values of control parameters like P, Z, v, d, or λ. If the P value of a rule system is close to 0.5, there is a tendency for chaotic trajectories to evolve; practically, this means trajectories with many different system states and a poor predictability of course. If we assume that most people prefer "well-ordered circumstances", i.e., clarity instead of the *Unübersichtlichkeit* proclaimed by Habermas, and that they need the certainty of knowing that circumstances will remain largely the same in the foreseeable future, so that they can rely on them, then it becomes immediately clear that most social systems, aside from revolutionary or other variable phases, have trajectories with simple attractors, and thus rule systems with high P values and low λ values, respectively. Strictly speaking, this has yet to be proven with real societies, but I think it is quite evident. Hence, if these hypotheses prove to be valid, a high P value would be a formal (social) universal in the strict Chomskyan sense, namely an (empirically proven) universal feature of social rule systems.

In any case, the control parameters can be interpreted as implicative formal universals, because all of these parameters obviously include the implication "if the control parameters of a social system have specific values, and if the rules of that system are kept constant, then a particular dynamics of that system will necessarily ensue". As we have seen, this implicative coherence is a mathematical property of rule systems, and thus an implicative formal universal of complex systems in general, according to Chomsky. Empirical analyses will have to find out whether the control parameters are not only implicative formal universals, but

also absolute ones: are there particular values appearing with a significant frequency in socio-cultural evolution? The conjecture of inequality obviously expresses such universal features: as most societies contain different degrees of inequality it follows that these societies merely generate rather simple dynamics; this is exactly what can be observed in reality. The conjecture of inequality expresses obviously such universal features: as most societies contain different degrees of inequality it follows that these societies generate only very simple dynamics, which in feet can be observed. Furthermore, the question of possible connections between substantial and formal universals arises: each of the universals listed by Brown corresponds in fact to a specific social rule. A systematic analysis of these rules might lead to the result that specific values of control parameters can be found in these rule systems. This is one of the fields - see below - where there is still a lot of research to be done in the future.[117]

Another example of a formal social universal seems to result from the game theoretical analyses of Axelrod as well as Nowak and May (see above). Hobbes' classical question why there is no permanent war of everyone against everyone could thus be answered in the following way: there is a formal universal b with an implicative nature, so that in the case of $b < 1.8$, cooperative behavior necessarily arises and becomes a social norm. However, as was shown in chapter 4., this is only valid if b itself is not disposable. The analyses in chapter 4. have also indicated that social systems are probably much more variable than languages.

We will now see that the same is true with regard to meta-parameters. Unfortunately, Brown did not pay any attention to the phenomenon of social change in his list, which may be partly due to the fact that ethnologists are usually only interested in reconstructing the actually valid rules of a given culture. Therefore, there seem to be no systematic analyses about the way and the degree in which social changes take place. Hence, the meta-parameters analyzed in section 3.3.3. can "only" be granted the character of implicative universals for now, according to the principle: if there are such and such values of the different meta-parameters, then the adaptive success, i.e., tempo and precision of approaching the goal values, will be such and such. Again, it is very likely that some of the meta-parameter values will prove to be absolute

[117] In the meantime, Jörn Schmidt and I formulated a general hypothesis about social dynamics, namely the "Theorem of Social Inequality". The theorem postulates that social inequality generates simple dynamics and vice versa (Klüver and Schmidt 1999). As social inequality is certainly a social universal, social order, i.e. simple dynamics, seems to be a mathematical consequence of this universal and in this way social order may be *explained* as a universal also.

formal universals: in fact, I have explicitly pointed to the fact that the r parameter, the "principle of the sacred cows", corresponds to a very realistic situation, namely the fact that social systems try to change as little as possible at once. From this aspect, it can be presumed that low values of r occur with a high statistical significance in reality, and in this sense they are candidates for absolute formal universals. The same is true, accordingly, for the problem of the dimensionality of spaces of social interactions.

The question that immediately ensues from the search for (formal and/or substantial) social universals is what can actually be "explained" with such results. In other words: will we find what has always been searched for as the universal laws of human history behind those social universals? Again, it is useful here to look at how linguistic universals are dealt with in linguistics.

Chomsky and his school are known to have taken the existence of linguistic universals as a proof for the fact that language has a biological basis, in that there have to be "innate (linguistic) ideas" (Chomsky) or "linguistic genes" (Pinker). Linguistics does not become reducible to biology by that, but it becomes part of a natural science that starts out from the neurophysiological processes in the brain, and then describes linguistic processes as results of the former with regard to the given social milieu, or peculiarities of the specific language. On the one hand, the linguistic universals, as general regularities, *refer* to the regularities of neurophysiological processes; and on the other hand, they *explain* linguistic processes from the neurophysiological regularities. One could say that universals are the manifestations of similarities of neurophysiological foundations within the great variety of distinct languages.

It may be that the social (substantial) universals found by Brown likewise refer to generic general and similar features, and that they, as "social competence" (Habermas), correspond to fixed synaptic links in the brain. One can certainly maintain the plausibility of such considerations without falling victim to a socio-biological reductionism; in fact, the argument is not that a specific behavior is genetically predetermined, but that there is a similar *mould* of the brain, which has to be understood as a product of both phylo- and ontogenetic learning processes, and which always bears the possibility of changing. As a result, the great variety of substantial social universals would point to the fact that a generic neural "basic equipment" rewards similar solutions, as long as this is tolerated by the circumstances; however, the same basic neural equipment makes it possible for the species to try other solutions, and thus change the neural equipment by force of new learning processes. So, this " basic neural equipment" may only be understood as

an adaptive, and thus, variable, system itself, and not as an ensemble of unchangeably predetermined genetic structures. Anyway, at the present time these are still speculations.

The mathematical properties of rule systems in the form of control and meta-parameters, by contrast, refer to characteristics of complex systems in general, and so they have got nothing to do with assumptions on genetic features. Thus, empirical analyses about the adaptive behavior of real social systems must always consider several possible explanatory levels: one possibility is that there are purely mathematical reasons for the behavior of a system, like in the case of the formal system SOZGEN (see above, section 4.1). As was described above, SOZGEN solved a double task "on its own", namely by annulling a part of its rules, after having attained particular goal values by rule variation. This guaranteed that the attained favorable state was only changed slightly for a relatively long period. The reasons for this behavior are completely due to particular mathematical properties of the formal system - it chose the mathematically simplest solution. This formal universal can be frequently observed empirically, as was already mentioned.

Another possibility is the occurrence of a behavior that cannot be explained by purely mathematical means, and that might even seem suboptimal in the light of a mathematical analysis. Such behavior may have to do with traditions, well-worn habits of thinking and behaving, and other constraints that can only be fully understood against the background of the particular history of the system. History then functions as a constraint, in complete accordance with Bateson: it narrows down the space of mathematically possible "favorable" behavior, so that possibly only suboptimal options remain; regardless of the fundamental variability of every social system, additional influences are necessary in such cases in order to widen the space of possibilities again.[118] But even in those cases one could discern universal patterns: adaptive systems are free to test the most favorable possibilities only in a mathematical sense. In reality one always has to consider which meta-parameter values are actually at the systems' disposal, and to what extent only a change of the system rules can lead to a realization of meta-parameter values that are relevant for the adaptation.

At the present time, these remarks are not (and cannot be) very concrete. However, we can already emphasize that formal social universals first of all refer "only" to the fact that social systems, like any complex adaptive system, are subject to specific mathematical

[118] This can be seen, e.g., from the recent history in the former Yugoslavia, which could obviously only stumble from one unfavorable possibility to next before the NATO intervention (written in 1997 before the Koshovo war).

regularities, which usually do not force unambiguous solutions, but instead characterize a space of possibilities. Each individual system looks for its own solutions within that space, which may be - but does not have to be - partly predetermined by its particular history. In other words: the concrete dynamics of a social system would have to be analyzed as a constraint by its concrete history of the space of possibilities, which in itself would have to be analyzed with regard to its self-produced constraints.

We may leave it open whether history is "the history of class struggles" (Marx), the change of forms of communication (Habermas), or perhaps something else. There probably are no such quasi-natural scientific laws, or at least none that are universal. So let us be content with a goal that is more modest as regards content, though more ambitious as regards methods: to gain insights into the behavior of social adaptive systems and thereby to find some mathematical characteristics in the form of formal universals. That might be a way towards a science of social reality that is theoretically and methodically consistent.

5.3. CONCLUSION AND PROSPECTS

> "Does it make sense to think about the human brain as a complex adaptive system" I asked. "That's right without doubt" she [Patricia Churchland] answered quickly. "But what is the use? Which research program follows from it?" (Lewin 1992, 205)

At the beginning of this work, I stressed that it would not make much sense at present to expect a new theory of social systems on the basis of the methodical approaches presented here. Instead, I named four aspects which may, and shall, serve as building blocks for such a future theory. Let us recall them once again to assess the level that has been reached with this work. As a result, we will encounter a number of advanced research questions, which could be turned into a research program, and will answer Churchland's question about the sensibility of conceiving social systems as adaptive dynamic systems.

(a) It is obviously possible - and certainly necessary as well - to define the basic concepts of dynamic systems as precisely as it would be required in a consistent science. Apparently, this does not only apply to the terminology of natural scientific system analyses, but also to the much more complex objects of the "cultural and cognitive sciences". Thus, it is no longer necessary for the latter to use the concepts of the natural sciences as metaphors in their own (social) contexts, with the inevitable notorious vagueness; instead, it is possible to formulate social systems theories with any accurateness that is desired. This particularly applies to those issues - rule variation, system learning, self-modelings,

self-enlargements, self-referentiality - that cannot be grasped by purely natural scientific concepts. For that purpose, one generally has to rid oneself of a more than heuristic orientation to the terminology of the natural sciences, and gain one's own precision from the place where any exact science has always received it: the formal terminology of mathematics and logics. In this sense (and not only in this), this is a very Kantian enterprise.

(b) Social phenomena of any complexity can be modeled formally on the basis of a terminology that is (re)formulated by this means. This point could only be exemplified with regard to the most important aspects of social systems in this study, but I hope that it has proven the plausibility of my claim. In somewhat aphoristic terms: if a social fact or process can be described precisely at all, then it can also be modeled formally by use of the terminology that was described here. It might of course be due to my lack of imagination, but presently I cannot see any case within the sociological literature which could not be mapped into formal models, and then be dealt with by analyzing these models. This assertion becomes strongly evident, in my opinion, by the heterogeneity of the examples that were presented in this study; if there are any counterexamples, I would be very curious to know them.

(c) Obviously, the analysis of "pure" formal systems, i.e., with no consideration for empirical aspects, can already provide important insights about the behavior of adaptive and thus also social systems. Of course, such experiments with formal systems cannot replace empirical research; but they can give empirical research a theoretical orientation by indicating which implicative formal universals may be transformable into absolutely valid substantial universals. It should be obvious by now that by analyzing formal systems, one can and must also gain the necessary knowledge of the formal apparatus.

Despite the great number of open questions of research, we can already obtain some informal insights. The most important one, in my opinion, is the greater precision in defining the relation between natural and social systems. This does not only concern the general realization that social systems cannot be described adequately by conceptual borrowings from the natural sciences, which has been mentioned again and again in this study; it also concerns an aspect that runs like a thread through the analysis of modern societies:

Ever since the introduction of the concept of *Entfremdung* (alienation) by Marx, there have been continual variations on a theme that is expressed most vividly in the famous quotation from Lukacs (1920/1971) about modern society being "second nature"; Adorno is known to have picked up this point with a different denotation. What was

thereby meant, in general terms, was that by "objective constraints" (*Sachzwänge*), "rationalizations", "objectivations", etc. modern societies are more and more experienced by their members like the inescapable laws of nature, to which they can only surrender, and which turn the individual's freedom and self-determination into idealistic fiction. Weber's "*Gehäuse der Hörigkeit*" (shell of obedience) belongs to this tradition, as well as the "*Entdifferenzierung von Lebenswelt und Systemen*" (dedifferentiation of life world and systems) in Habermas; the "*Risikogesellschaft*" (risk society) (Beck 1986), a model of society that deals with the coercion of the self-produced risks of large-scale technology projects, can easily be understood as a new variant of this classical concept.

The basic conception of social systems as adaptive, and thus variable, systems that was developed here paints a completely different picture. It should be out of the question that the variability of societies, and thus their adaptive potential, have been accomplished in the most radical form in modern societies. To vary a famous quotation from Marx: only in modern societies is it actually possible to give a description of society that is based on its variability. We have seen again and again that, and to what extent, the variability of social systems in its various facets differs fundamentally from that of natural systems; the more variable a system, the less "natural" it is. In other words: modern societies are no "second nature," but something completely different *in principle*; on the other hand, the less modern a society is, the more adequately it can be described with the instruments of the natural scientific systems theories, and the more it is a reproduction of natural circumstances by social means.

The characteristics of modern societies that were analyzed by Marx ,Weber, Adorno, Habermas, and others do of course exist. But rather than being the inescapable fate of a second nature, they are aspects of social systems which have originally produced those phenomena by adaptive rule variations, and which also tackle them by further rule variations. The obvious capability of modern variable societies to manage their self-produced ecological problems, at least in a better way than previous societies, is only *one* indication of this. To emphasize it once again: the adaptive variability of modern societies is no guarantee for a timely and sufficient solution to all problems. Meticulous analyses will be necessary, which require a combination of empirical research and computer modelings, in order to evaluate whether the variation potential of modern societies is actually sufficient to manage the great number of problems. Therefore, there is no cause for cheerful optimism at all. Nevertheless, the history of modern (capitalist) societies is mainly the history of their continual changes, and not the history of their

inescapable coercions. One only has to look at the history of the numerous analyses that proved the unavoidable end of capitalism. It cannot be assumed that this will change.

Besides, this throws a distinct light on the great quarrels of world politics in this century: it is very conspicuous that in any of the great "hot" or "cold" conflicts, the Western bourgeois-capitalist societies always won through - against their semi-feudal, fascist, or "real socialist" opponents. Unless one wants to postulate universal strategic laws at work to explain this phenomenon, as has been done, e.g., by the historian Kennedy (1987), our analyses and results can offer a relatively simple explanation: the Western societies obviously had, and have, the greatest potential of variability, and thus the greatest adaptive success. The fact that such factors can be decisive was shown, e.g., by the results of SOCVAR in chapter 4: the societies with the most favorable meta-parameter values always succeeded. It is an open question whether Western societies are morally superior to other societies, which not only Habermas thinks; on the other hand, I have no doubt that they are superior with regard to their adaptivity, i.e., they have the highest variability. It is the fundamental social self-conception of these societies that practically forces their variability; dictatorships of whatever provenance, by contrast, must insist on the greatest possible stagnation of the rules, i.e., on a low variability, with the same necessity.

Hence, a reanalysis of the well-known classifications of different social formations would suggest itself in view of the question whether particular social forms correspond to equally particular variability degrees. The cursory reflections in section 3.3 about the relation between variability and complexity suggest such an assumption: the more variable a society is, the more different states can it attain (algorithmic complexity), and the longer the description of its hybrid rule system (effective complexity) has to be; the more complex a society is (has become), the greater is the number of variations it must have carried out since its simple initial states, and the higher is its variability as a consequence - provided that it has not blocked its variability on its own, like the feudal Chinese or Japanese systems. A lot of basic research is left to be done here as well. Anyway, on the basis of our experiments' results, it seems necessary to me to take a society's variability - measured by its meta-parameters - as the criterion, and as an explanation, of its actual evolutionary success.

The extreme difference between the developments of medieval Europe and the feudal Islamic, Chinese, and Japanese advanced civilizations, which were at least at an equal developmental level at that time, could then be explained following an idea by Needham (1970):

heterogeneous Europe brought its distinct societies into constant competition with each other, which forced them to adapt continually to their more successful competitors. In other words: Europe developed as an advanced feudal civilization with distinct societies, each of which had a high degree of variability. So, *from the point of view of systems theory*, its further development was almost inevitable, because recombinatory innovations and self-enlargements had to be activated as a potential of every social system. A more recent example seems to be the similarly extreme difference between the developments of the newly industrializing countries in South East Asia and most African countries: it is virtually impossible to overlook their different degrees of variability.

It is evident that different degrees of variability "per se" do not explain anything, because the *sociological* question must of course explore the reasons for these differences. However, these are empirical questions, which means that they are historical in the classical sense. Studying the degrees of variability offers the possibility of defining the evolutionary success or failure of different social systems precisely, and thus of dealing with the examination of the reasons for specific evolutionary processes, and the subsequent question about their probable future, in a rather exact way.

Social sciences are not least sciences of action, and so one might expect at the end of a study like this some remarks on how these systems might be used with regard to practical interventions. However, I cannot share the frequently expressed optimistic assertion of there already being successful systemic approaches of intervention and control (e.g. Willke 1995). Our knowledge about the dynamics of social systems is still much too indefinite, and we are far from having an elaborate theory of social systems - a fact that has been stressed throughout this work; therefore, it would be dishonest to pretend that there already are easily graspable answers. First of all we need a better, that is to say, a more detailed understanding. As a consequence, I can only recommend *one* thing, and that is to test the probable, or merely possible, consequences of any intervention by precise model building and simulations before really intervening into a complex system. In other words: apply the methodical procedure described in this study.

Of course, we can already draw some conclusions from our current results that indicate aspects that should be considered in any attempts to intervene or control. For example, in the case of a social system with a low variability, one should always test which rule variation would provide better trajectories for the system. Because of the system's low variability, we may then presuppose that it really attains those trajectories. If, by contrast, we have a system with a high variability, then we would certainly not be very successful if we tried to influence it by

changing single states or rules: in all probability, the system would undo the externally affected changes, or transform them into totally different system states and rules, simply because of its high variability. In this case, the principle of "contextual" control as recommended by Willke (1995) suggests itself: change the environment of the system in such a way that the system itself has to use its variability in order to fulfill its environmental requirements. We have constantly applied just this form of system "control" in our hybrid systems, because even their simple variants cannot anymore be controlled by simple interventions.

So, presently our results do not lead to a control "theory", but at least they can reveal concrete prospects of how such a theory might be developed. However, this requires the construction of complex programs and experiments, more experiments, and experiments with them once again. Unfortunately, there is no easier way towards a social control theory, but that is just the way it is in all ambitious fields of knowledge. Anyhow, such a control theory is absolutely necessary, because there can definitely be no question of a self-established harmony in the case of such complex, that is, with adaptive-variable systems. The conception of the "invisible hand" that resolves everything by itself to everyone's satisfaction, which is conjured up again and again in neoliberalist economic concepts, must seem almost ridiculously naive in the light of our experiments. Once again we can refer to Bateson here, who has provided a thorough and scathing critique of those laissez-faire conceptions (1979).

We started this last subchapter with Patricia Churchland's question about a research program; I fully agree with her that this is the only way to give any theoretical methodical approach a practical meaning. Unfortunately (?), it is quite easy to formulate such a program; one only has to list all the questions that had, for various reasons, to be left open in this study. We can basically specify two complexes:

1) On a formal level, we are definitely far from having sufficient knowledge about the behavior of formal self-organizing and adaptive systems. Though we know a number of control parameters, and have at least sound assumptions on some meta-parameters, we still know little or nothing about the mathematical connections between the single control parameters, the single meta-parameters, and between control and meta-parameters; the "mathematics of self-referentiality", to use the term I assigned to that problem area in chapter 3, is still in its infancy, in spite of all the results achieved so far. Another thing that is largely unknown is the actual dynamics of hybrid systems in dependence on the different meta-parameters: we studied the effects of the meta-parameters on the adaptive success of hybrid systems, which is certainly sensible; however,

we also realized in our experiments with SOZGEN - the program that can turn its meta system on or off - that it is necessary to analyze the trajectories of such systems as well. As I said above, SOZGEN attained the additional goal of staying close to its achieved favorable state values after the GA was turned off by simply turning off most of its base rules. By that, it stayed close to a simple attractor, though not every time. There are of course other solutions for this problem as well, which would have to be investigated; one important point is that real systems are usually unable to carry out this simple kind of solution. The trajectories of TRISOC (4.4.2.) show on an advanced level how a high sensibility of the system is caused both by its variability and its dependence on rather different optimization goals at the same time; systems of that kind can practically never rest, no matter how much this might be wished for by the people (especially by politicians of course).

After all, the trajectories of hybrid or meta dynamic systems result from the "partial trajectories" which make up the dynamics of the base system *between* the rule variations. From that arises a question that is completely open as yet, namely how an "overall trajectory" can be described as a combination of partial trajectories, in dependence on specific control and meta-parameters. Presently, it is still rather unclear how this might be achieved, and which statements could be derived therefrom about general properties of hybrid systems, though our experiments with systems like SOZION, SOZGEN, and TRISOC can already give some information about this. The list of problems could be increased.

2) An even more fundamental desideratum is of course the question of how to fill the concepts and resultant models that were presented here empirically. Each of our models was well explained and justified with reference to well-known sociological subjects, but of course that does not necessarily mean that they are empirically valid in the usual sense. From an empirical point of view, this study consists of plausibility arguments, which definitely make sense and are indispensable here; however, the actual empirical validation is still to come. In particular, we know that the results from simple systems with relatively few elements and rules are not necessarily transferable to systems that are much greater in quantity, which is usually the case with real systems - at least if one wants to model complete societies. To achieve this, a lot of single facts must be taken into account, and the use of mainframe computers with much greater capacities than presently the norm in the social sciences is required.

This work can of course not be done by a relatively small research group like ours at the University of Essen alone. An approach like this takes sociology to a stage where it has to organize its research activities

in units larger than single groups - in graduate programs, collaborative research centers, or special institutes, as has been a common practice in empirical social research for a long time. A research program like this requires a standing interdependence of theory formation, preparation of facts, model building, and computer experiments. These remarks on the organization of research are not intended to argue away this study's obvious desiderata. I rather wanted to point to the fact that methodical approaches orientated to the natural sciences need appropriate research organizations in order to be practically feasible.

Regardless of these remarks on sociological research, it is quite obvious which object of research should be tackled next - by us and hopefully by others too: we will necessarily have to deal with the exemplary preparation of historical and current social processes, take all available data into account, and then transform them into appropriate computer models and study them experimentally in the computer. This will hopefully lead to a substantiation of the general concepts that have been developed in this study, and also to new insights into the empirical constraints and, possibly, substantial universals that underlie real social systems. The goal is clear, and so is the way to reach it - now let us try to follow it.

On the condition that these and many other questions are answered, and the goal of a sociology that is strict, consistent, *and* rich in content is within reach - will we realize that we, as social individuals, are only objects of universal mathematical regularities? In other words: will the nightmare of the inescapable second nature repeat itself on a higher level, by the realization that any variability of social systems is determined by the logic of self-referential mathematical rules (including biological constraints), and that our creative freedom to influence history is just our imagination? I think not. Our experiments with the rather simple systems that were described here have already demonstrated that the space of freedom of hybrid systems is usually much too great to allow talk of unambiguous determinations. Thus, we are dealing with a dialectics, to use a classical philosophical term, between mathematical constraints and the freedom spaces *included in them*. This point was already expressed in an unsurpassed way 500 years ago by one of the first European social scientists, Nicolo Machiavelli, and I can take his words as an ending:

> "God is not willing to do everything, and thus take away our free will and that part of glory which rightly belongs to us."[119]

[119] The Prince (Il Principe), my translation from the German version "Der Fürst", J.K.

REFERENCES

Andersen, E.S. (1992) Complexity and Language Acquisition: Influences on the Development of Morphological Systems in Children, in J.A. Hawkins and M. Gell-Mann (eds.), *The Evolution of Human Languages,* Addison Wesley, Reading (Mass.), pp. 241-272

Anderson, P.W., Arrow, K., and Pines, D. (1989)*The Economy as an Evolving Complex System,* Addison Wesley, Reading (Mass.)

Aoun, J. (1994) A Brief Representation of the Generative Enterprise, in J.A. Hawkins and M. Gell-Mann (eds.), *The Evolution of Human Languages,* Addison Wesley, Reading (Mass.), pp. 121-136

Apel, K.O. (1979) *Die Erklären-Verstehen-Kontroverse in transzendental-pragmatischer Sicht*, Suhrkamp, Frankfurt

Axelrod. R. (1984) *The Evolution of Cooperation,* Basic Books, New York

Axelrod, R. (1987) The Evolution of Strategies in the Iterated Prisoner`s Dilemma, in L. Davis (ed.), *Genetic Algorithms and Simulated Annealing*, Morgan Kauffman, Los Altos, pp. 32-41

Bahr, D.B. and Passerini, E. (1998), Statistical Mechanics of Opinion Formation and Collective Behavior: Micro-Sociology, *Journal of Mathematical Sociology 23*, 1-49

Banks, D.L. and Carley, K.M. (1997) Models for Network Evolution, in P. Doreian and F. Stokman (eds.), *Evolution of Social Networks*, Gordon and Breach, Amsterdam, pp. 209-232

Bar-Yam, Y. (1997) *Dynamics of Complex Systems.* Addison Wesley, Reading (Mass.)

Barrow, J. D. (1994) *Theorien für Alles. Die Suche nach der Weltformel,* Rowohlt, Reinbek

Bateson, G. (1972) *Steps to an Ecology of Mind,* Chandler, London

Bateson, G. (1979) *Mind and Nature. A Necessary Unity,* Chandler, London

Beck, U. (1986) *Risikogesellschaft,* Suhrkamp, Frankfurt

Bedau, M.A. and Packard, N. (1992) Measurement of Evolutionary Activity, Teleology and Life, in C.G.Langton, C. Taylor, J.D. Farmer, and S. Rasmussen, (eds.), *Artificial Life II,* Addison Wesley, Reading (Mass.), pp. 431-461

Beer; S. (1974) *Designing Freedom,* Wiley, London

Belew, R., McInerney, J., and Schrandolph, N. (1992) Evolving Networks: Using the Genetic Algorithm with Connectionist Learning, in C.G.Langton, C. Taylor, J.D. Farmer, and S. Rasmussen (eds.),*Artificial Life II,* Addison Wesley, Reading (Mass.), pp. 511-547

Beltrami, E. (1993) *Mathematical Models in the Social and Biological Sciences. Jones and Bartlett Publishers*

Ben-David, J. (1971) *The Scientist`s Role in Society.* Englewood Cliff (N.J.)

Bennett, C.H. (1988) Logical Depth and Physical Complexity, in R. Herken (ed.), *The Universal Turing Machine. A half-century Survey,* Kammerer und Unverzagt, Hamburg, pp. 227-257

Berger, P. and Luckmann, T. (1966) *The Social Construction of Reality,* Doubleday, New York

Berlekamp, E., Conway, J.H., and Guy, R. (1982) *Winning Ways for Your Mathematical Plays,* Academic Press, New York

Bersini, H. (1991) Immune Network and Adaptive Control, in P. Bourgine, and F.J.Varela (eds.), *Towards a Practice of Autonomous systems. Proceedings of the First European Conference on Artificial Life,* MIT Press, Cambridge (Mass.), pp. 217-226

Boerlijst, M. and Hogeweg, P. (1992) Self-Structuring and Selection: Spiral Waves as a Substrate for Prebiotic Evolution, in C.G. Langton, C. Taylor, J.D. Farmer, and S. Rasmussen (eds.), *Artificial Life II,* Addison Wesley, Reading (Mass.), pp. 255-276

Böhme, G., van den Daele, W., and Krohn, W. (1977) *Experimentelle Philosophie,* Suhrkamp, Frankfurt

Bossel, H. (1992) *Modellbildung und Simulation,* Vieweg, Braunschweig,Wiesbaden

Bourdieu, P. Passeron, J.C. (1964) *Les Héritiers. Les Etudiants et la Culture,* Edition de Minuit, Paris

Bourdieu; P. (1976) *Entwurf einer Theorie der Praxis,* Suhrkamp, Frankfurt

Bourdieu, P. (1989) *Choses dites,* Edition de Minuit, Paris:. German: (1992) *Rede und Antwort,* Suhrkamp, Frankfurt

Bourgine, P. and Varela, F.J. (eds.) (1991) *Towards a Practice of Autonomous Systems. Proceedings of the First European Conference on Artificial Life,* MIT Press, Cambridge (Mass.)

Bourgine, P. and Varela, F.J. (1991 a) Introduction: Towards a Practice of Autonomous Systems, in P. Bourgine and F.J. Varela.(eds.), *Towards a Practice of Autonomous systems. Proceedings of the First European Conference on Artificial Life,* MIT Press, Cambridge (Mass.)

Brause (1991) *Neuronale Netze,* Teubner, Stuttgart

Briggs, J. and Peat, F.D. (1993) *Turbulent Mirrow,* Harper and Row, New York.

Brown, D.E (1991) *Human Universals,* Temple University Press, Philadelphia

Cariani, P.(1992) Emergence and Artificial Life, in C.G.Langton, C. Taylor, J.D. Farmer, S. Rasmussen (eds.), *Artificial Life II,*Addison Wesley, Reading (Mass.), pp. 775-797

Carley, K.M. (1997) Organizations and Constrint - Based Adaption, in R.A. Eve, S. Horsfall and M.E. Lee, *Chaos, Complexity, and Sociology. Myths, Models, and Theories,* Sage, London, pp. 229-242

Casti, J. (1997) *Would-be worlds: How simulation is changing the frontiers of science,* Wiley, New York

Chaitin, G.J. (1987) *Information Randomness. Papers on Algorithmic Information Theory,*World Scientific, Singapore

Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge (MA). Deutsch: (1969) *Aspekte der Syntax-Theorie,* Suhrkamp, Frankfurt

Churchland, P. and Sejnowski, T. (1992) *The Computational Brain*, MIT Press, Cambridge, London

Comrie, B. (1992) Before Complexity, in J.A. Hawkins and M. Gell-Mann (eds.), *The Evolution of Human Languages,* Addison Wesley, Reading (Mass.), pp. 193-212

Conte, R. and Castelfranchi, C. (1994) Mind is not enough, in N. Gilbert and J. Doran (eds.), *Simulating Societies. The Computer Simulation of Social Phenomena,* UCL Press, London, pp. 267-286

Coveney, P. and Highfield, R. (1990) *The Arrow of Time*, Allen, London

Crutchfield, J.P. and Young, K. (1990) Computation at the Onset of Chaos, in W. H. Zurek (ed.), *Complexity, Entropy and Physics of Information,* Addison Wesley, Redwood City, pp. 223-270

Daele, van den W. (1977) Die soziale Konstruktion der Wissenschaft - Institutionalisierung und Definition der positiven Wissenschaft in der 2. Hälfte des 17. Jahrhunderts, in G. Böhme, W. Daele, van den, and W. Krohn (eds.), *Experimentelle Philosophie,* Suhrkamp, Frankfurt, pp. 129-182

Davis, J.A. (1967) Clustering and structural balance in graphs. *Human Relations 20*, 181-187

Davis, L. (ed.), (1987) *Genetic Algorithms and Simulated Annealing,* Morgan Kauffman, Los Altos

Dawkins, R. (1976) *The Selfish Gene,* Oxford University Press, Oxford

Dawkins, R. (1987) *The Blind Watchmaker,* Norton, New York

Dewdney, A.K. (1988 a) Life in drei Dimensionen, *Computer-Kurzweil II., Sonderheft Spektrum der Wissenschaft*

Dewdney, A.K. (1988) Rechnen mit Spaghetti, *Computer-Kurzweil II., Sonderheft Spektrum der Wissenschaft*

Dewdney, A.K. (1990) Der Mandelbus, *Computer-Kurzweil IV., Sonderheft Spektrum der Wissenschaft*

Dörner, D. (1992) *Die Logik des Mißlingens,* Rowohlt, Reinbek

Doreian, P. and Stokman, F. (1997) The Dynamics and Evolution of Social Networks, in P. Doreian and F. Stokman (eds.), *Evolution of Social Networks,* Gordon and Breach, Amsterdam, pp. 1-17

Dreyfus, H.L., Dreyfus, S.E. (1987) *Künstliche Intelligenz,* Rowohlt, Reinbek

Dupuy, J.P. and Varela, F.J. (1992) Understanding Origins. An Introduction, in F.J. Varela and J.P. Dupuy (eds.), *Understanding Origins,* Kluwer Academic Publishers, Dordrecht, pp. 1-25

Edelman, G.M. (1992) *Bright Air, Brilliant Fire - On the Matter of the Mind,* Basic Books, New York

Eder, K. (1976) *Die Entstehung staatlich organisierter Gesellschaften.* Piper, Frankfurt

Eigen, M. and Schuster, P. (1979) *The Hypercycle: A Principle of Natural Self-Organization,* Springer, Berlin

Epstein, J.M. (1997): *Nonlinear Dynamics, Mathematical Biology and Social Science*, Addison Wesley, Redwood

Esser, H. (1991) Die Rationalität des Alltagshandelns. Alfred Schütz und "Rational Choice", in H. Esser and K.G. Troitzsch (Hrsg.), *Modellierung sozialer Prozesse,* Informationszentrum Sozialwissenschaften, Bonn, pp. 235-282

Esser, H. and Troitzsch, K.G. (eds.), (1991) *Modellierung sozialer Prozesse,* Informationszentrum Sozialwissenschaften, Bonn:

Eve, R.A., Horsfall, S., and Lee, M.E. (1997) *Chaos, Complexity, and Sociology. Myths, Models, and Theories,* Sage, London

Farmer, J.D. (199O) A Rosetta Stone for Connectionism, in Forrest, S. (ed.), Emergent Computation, *Physica* ***D 42,*** 1-31

Fararo, T. (1989) *The Meaning of General Theoretical Sociology. Tradition and Formalization,* Cambridge University Press, Cambridge

Fararo, T. (1997) *Reflections on Mathematical Sociology,* in *Sociological Forum 12*, 73-101

Fogel, D.B. (1993) Evolving Behaviors in the Iterated Prisoner`s Dilemma, in *Evolutionary Computation* ***1,*** 77-97

Forrest, S. (ed.), (1990) Emergent Computation, *Physica* ***D 42,*** 1-3, North-Holland, Amsterdam

Freeman, D. (1983) *Margret Mead and Samoa: The Making and Unmaking of an Anthropological Myth,* Cambridge University Press, Cambridge (Mass.)

Freeman, L. (ed.), (1989) *Research Methods in Social Network Analysis,* George Mason University Press, Fairfax

Freeman, L. (1989 a) Social Networks and the Structure Experiment, in L. Freeman (ed.), *Research Methods in Social Network Analysis,* George Mason University Press, Fairfax, pp. 11-40

Geertz, C. (1972) Deep Play. Notes on the Balinese Cockfight, in *Daedalus* **101,** 1-37

Gelfand, A.E. and Walker, C.C. (1984) *Ensemble Modeling,* Marcel Dekker, New York

Gell-Mann, M. (1992) Complexity and Complex Adaptive Systems, in J.A. Hawkins and M. Gell-Mann (eds.), *The Evolution of Human Languages,* Addison Wesley, Reading (Mass.), pp. 3-18

Gell-Mann, M. (1994) *The Quark and the Jaguar*, Freeman, New York

Gerhard, P. and Schuster, H. (1995) *Das digitale Universum.* Vieweg, Braunschweig,Wiesbaden

Giddens, A. (1984) *The Constitution of Society. Outline of the Theory of Structuration,* Polity Press, Cambridge

Giesen, B. (1980) *Makrosoziologie. Eine evolutionstheoretische Einführung,* Hoffmann und Campe, Hamburg

Gilbert, N. and Doran, J. (eds.), (1994) *Simulating Societies. The Computer Simulation of Social Phenomena.* UCL Press, London

Gilbert, N. (1995) Emergence in Social Simulation, in N. Gilbert and R. Conte (eds.), *Artificial Societies. The Computer Simulation of Social Life,* UCL Press, London, pp. 144-156

Gilbert, N.and Conte, R. (eds.), (1995) *Artificial Societies. The Computer Simulation of Social Life,* UCL Press, London

Gleick, J. (1987) *Chaos: Making a New Science.* Penguin,New York

Goldberg, D.E. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning,* Addison Wesley, Reading (MA)

Goodall, J. (1986) *The Chimpanzees of Gombe. Patterns of Behavior,* Belknap Press of Harvard University Press, Cambridge (Mass.)

Goonatilake, S. and Khebbal, S. (eds.), (1995)*Intelligent Hybrid Systems.* John Wiley & Sons, London

Goonatilake, S. and Khebbal, S. (1995 a) Intelligent Hybrid Systems. Issues, Classifications and Future Directions, in Goonatilake S. and S. Khebbal (eds.), *Intelligent Hybrid Systems,* John Wiley Sons, London

Gould, S. J. and Lewontin, R.L. (1979) The Spandrels of San Marco and the Panglossian Paradigm, *Proc. Royal Society London* ***B 205,*** 581-598

Grewendorf, G., Hamm, F.,and Sternefeld, W. (1987),*Sprachliches Wissen,* Suhrkamp Frankfurt

Grossberg, S. (1982) *Studies of Mind and Brain,* Reidel, Dordrecht

Gumermann, G.J. and Gell-Mann, M. (1994) (eds.), *Understanding Complexity in the Prehistoric Southwest,* Addison Wesley, Reading (Mass.)

Habermas, J. (1968) *Technik und Wissenschaft als "Ideologie",* Suhrkamp, Frankfurt

Habermas, J. (1976) Moralentwicklung und Ich-Identität, in J. Habermas,*Zur Rekonstruktion des Historischen Materialismus,* Suhrkamp, Frankfurt, pp. 63-91

Habermas, J.(1981) *Theorie des kommunikativen Handelns I und II,* Suhrkamp, Frankfurt

Harnad, S. (1994) Artificial Life: Synthetic vs. Virtual, in C.G. Langton (ed.), *Artificial Life III,* Addison Wesley, Reading (Mass.), pp. 539-552

Hawking, S.W. (1988) *A Brief History of Time. From the Big Bang to Black Holes,* Bantam Books, New York

Hawkins, J.A. and Gell-Mann, M. (eds.), (1992) *The Evolution of Human Languages,* Addison Wesley, Reading (Mass.)

Hebb, D.O. (1949) *The Organization of Behavior,* Wiley, New York

Hegselmann, R. (1996) Cellular Automata in the Social Sciences. Perspectives, Restrictions and Artefacts, in R. Hegselmann, U. Mueller, and K.G. Troitzsch (eds.), *Modeling and Simulation in the Social Sciences from the Philosophy of Science Point of View,* Kluwer Academic Publishers, Dordrecht, pp. 209-234

Hegselmann, R., Mueller, U., and Troitzsch, K.G. (eds.), (1996)*Modeling and Simulation in the Social Scienes from the Philosophy of Science Point of View,* Kluwer Academic Publishers, Dordrecht

Heiden, U. an der (1992) Selbstorganisation in dynamischen Systemen, in W. Krohn and G. Küppers (eds), *Emergenz: Die Entstehung von Ordnung, Organisation und Bedeutung,* Suhrkamp, Frankfurt, pp. 57-88

Helm, G. (1991) *Symbolische und konnektionistische Methoden der menschlichen Informationsverarbeitung,* Springer, Berlin

Hillis, D. (1992) Co-Evolving Parasites Improve Simulated Evolution as an Optimization Procedure, in C.G. Langton, C. Taylor, J.D. Farmer, S. Rasmussen (eds.), *Artificial Life II,* Addison Wesley, Reading (Mass.), pp. 313-324

Hofstadter, D.R. (1985) *Gödel, Escher, Bach an Eternal Golden Braid,* Basic Books, New York

Hofstadter, D.R. (1986) Mathematical Chaos and Strange Attractors, in D.R. Hofstadter, *Metamagical Themas,* Penguin, Harmondsworth, pp. 364-395

Hofstadter, D.R. (1986*) Metamagical Themas*, Penguin, Harmondsworth

Hofstadter, D.R. and Dennett, D.C. (1987) Reflexionen (zum Artikel von Searle), in D.R. Hofstadter and D.C. Dennett (eds.),*Einsicht ins Ich ("The Mind's!" 1981),* Klett, Stuttgart, pp. 357-366

Holland, J.H, Holyoak, K.J., Nisbett, R.E., and Thagard, P.R. (1986)*Induction,* MIT Press, Cambridge (Mass.)

Holland, J.R. (1975) *Adaptation in Natural and Artificial Systems,* MIT Press, Ann Arbor: University of Michigan Press

Holland, J.R. (1995) *Hidden Order.* Addison Wesley, Reading (Mass.)

Holland, J.R. (1998) *Emergence. From Chaos to Order,* Addison-Wesley, Reading (Mass.)

Hondrich, K.O. (1987) Die andere Seite sozialer Differenzierung, in H. Haferkamp and M. Schmid (eds.), *Sinn, Kommunikation und soziale Differenzierung,* Suhrkamp, Frankfurt, pp. 275-303

Hopfield, J.J. (1982) Neural Networks and Physical Systems with Emergent Collective Computational Abilities, *PNAS 79,* 2554-2558

Horgan, J.(1995) Komplexität in der Krise, *Spektrum der Wissenschaft, September 1995*, pp. 58-64

Kari, J. (1992) Properties of Limit Sets of Cellular Automata, in N. Boccara, E. Goles, S. Martinez and P. Picco (eds.),*Cellular Automata and Cooperative Systems,* Kluver Academic Publishers, Dordrecht, pp. 311-321

Kauffman, S.A. (1992) Origins of Order in Evolution: Self-Organization and Selection, in F.J. Varela and J.P. Dupuy (eds.), *Understanding Origins. Contemporary Views on the Origin of Life, Mind and Society,* Kluwer Academic Publishers, Dordrecht, pp. 153-181

Kauffman, S.A. (1993) *The Origins of Order,* Oxford University Press, Oxford,

Kauffman, S.A. (1995) *At Home in the Universe,* Oxford University Press, New York

Kennedy, P. (1987) *The Rise and Fall of the Great Powers,* Random House, New York

Klemm, E. and Wittek, R. (1989) Codierung und Gruppierung von Netzwerkdaten, in T. Schweizer (ed.), *Netzwerkanalyse: Ethnologische Perspektiven,* Reimer, Berlin, pp. 93-111

Klüver, J. (1995) *Soziologie als Computerexperiment,* Vieweg, Braunschweig, Wiesbaden

Klüver, J. (1996) Simulations of Selforganizing Social Systems, in F. Faulbaum and W. Bandilla (eds.), *SoftStat 95, Advances in Statistical Software,* Lucius, Stuttgart, pp. 425-432

Klüver, J., Stoica, C., and Schmidt, J. (1997) Simulations of the Self Modeling of Social Systems by a Hybrid System, in F. Faulbaum and W. Bandilla (eds.), *SoftStat 97, Advances in Statistical Software,* Lucius, Stuttgart, pp. 411-417

Klüver, J. and Schmidt, J. (1999) *Control Parameters in Boolean Networks and Cellular Automata Revisited: From a Logical and a Sociological Point of View*, to be published in Complexity

Kneer, G. and Nassehi, A. (1994) *Niklas Luhmanns Theorie sozialer Systeme,* W. Fink, München

Knorr-Cetina, K. (1981) Introduction: The micro-sociological challenge of macro-sociology: towards a reconstruction of social theory and methodology, in K. Knorr-Cetina and A.V. Cicourel (eds.), *Advances in Social Theory and Methodology,* Routledge and Kegan Paul, Boston, London, pp. 1-48

Koza, J. R. (1992) *Genetic Programming,* MIT Press, Cambridge (Mass.)

Koza, J. R. (1994) Artificial Life: Spontaneous Emergence of Self-Replicating and Evolutionary Self-Improving Computer Programs, in C.G. Langton (ed.), *Artificial Life III,* Addison Wesley, Reading (Mass.), pp. 225-262

Krohn, W. and Küppers, G. (eds.), (1992) *Emergenz: Die Entstehung von Ordnung, Organisation und Bedeutung,* Suhrkamp, Frankfurt

Krohn, W. and Küppers, G. (1992 a) Zur Emergenz systemspezifischer Leistungen, in W. Krohn and G. Küppers (eds.), *Emergenz: Die Entstehung von Ordnung, Organisation und Bedeutung,* Suhrkamp, Frankfurt, pp. 161-188

Küppers, G. and Krohn,W. (1992) Selbstorganisation. Zum Stand einer Theorie in den Wissenschaften, in W. Krohn an G. Küppers (eds), *Emergenz: Die Entstehung von Ordnung, Organisation und Bedeutung,* Suhrkamp, Frankfurt, pp. 7-26

Lakatos, I. (1971) Falsification and the Methodology of Scientific Research Programmes, in I. Lakatos and A. Musgrave (eds), *Criticism and the Growth of Knowledge,* Cambridge University Press, Cambridge, pp. 91-196

Lakatos, I. and Musgrave, A. (eds.) (1971) *Criticism and the Growth of Knowledge,* Cambridge University Press, Cambridge

Langton, C.G. (1984) Self-Reproduction in Cellular Automata, *Physica 10 D,* 135-144

Langton, C. G. (ed.), (1988) *Artificial Life,* Cambridge University Press, Reading (Mass.)

Langton, C.G. (1988) Preface, in C.G. Langton (ed.), *Artificial Life,* Cambridge University Press, Reading (Mass.), pp. i-xxv

Langton, C. G. (1992) Life at the Edge of Chaos, in C.G. Langton , C. Taylor, J.D. Farmer, and S. Rasmussen (eds.), *Artificial Life II,* Addison Wesley, Reading (Mass.), pp. 41-91

Langton, C.G., Taylor, C., Farmer, J.D., and Rasmussen, S. (eds.) (1992) *Artificial Life II,* Addison Wesley, Reading (Mass.)

Langton, C.G. (ed.) (1994) *Artificial Life III,* Addison Wesley, Reading (Mass.)

Lansing, J.S. and Kremer, J.N. (1994) Emergent Properties of Balinese Water Temple Networks: Coadaptation on a Rugged Fitness Landscape, in C.G. Langton (ed.), *Artificial Life III,* Addison Wesley, Reading (Mass.), pp. 201-223

Leinfellner, W. (1995) Soziale Intelligenz und Rationalität: Konfliktlösen als Evolution, *Zeitschrift für Wisenschaftsforschung* ***9/10,*** 73-107

Levy, S. (1993) *Artificial Life. The Quest for a New Creation,* Penguin, London

Lewin, R. (1992) *Complexity. Life at the Edge of Chaos,* Macmillan, New York. German: *Die Komplexitätstheorie. Wissenschaft nach der Chaosforschung,* Hoffman und Campe, Hamburg (1992)

Lindgren, K. (1992) Evolutionary Phenomena in Simple Dynamics. in C.G. Langton , C. Taylor, J.D. Farmer, and S. Rasmussen (eds.),*Artificial Life II,* Addison Wesley, Reading (Mass.), pp. 295-312

Lucakcs, G. (1971) *Die Theorie des Romans. Ein geschichtsphilosophischer Versuch über die Formen der großen Epik,* Luchterhand, Neuwied

Luhmann, N. (1984) *Soziale Systeme,* Suhrkamp, Frankfurt

Luhmann, N. (1985) Die Autopoiesis des Bewußtseins, *Soziale Welt* ***36,*** 402-446

Luhmann, N. (1986) *Ökologische Kommunikation,* Leske & Buderich, Opladen

Lurija, A.R. (1993) *Romantische Wissenschaft,* Rowohlt, Reinbek

Malinowski, B. (1944) *A Scientific Theory of Culture*, Chapel Hill

Mandelbrot, B. (1982): *The Fractal Geometry of Nature*, Freeman, New York

Maturana, H.R. (1982) *Erkennen: Die Organisation und Verkörperung von Wirklichkeit,* Braun-schweig,Wiesbaden

Maturana, H. R., Varela, F.J., and Uribe, R. (1982) Autopoiese: Die Organisation lebender Systeme, ihre nähere Bestimmung und ein Modell, in H.R. Maturana *Erkennen: Die Organisation und Verkörperung von Wirklichkeit,* Vieweg, Braunschweig, Wiesbaden, pp. 157-169

May, R.M. (1976) Simple Mathematical Models with Very Complicated Dynamics, *Nature* ***261,*** 459

Maynard Smith, J. (1974) *Models in Ecology,* Cambridge University Press, Cambridge

Maynard Smith, J. (1982) *Evolution and the Theory of Games,* Cambridge University Press, Cambridge

Maynard Smith, J. (1993) *The Theory of Evolution,* Cambridge University Press, Cambridge

Mayntz, R. (1988) Funktionelle Teilsysteme in der Theorie sozialer Differenzierung, in R. Mayntz, R. Rosewitz, U. Schimank, and R. Stichweh, *Differenzierung und Verselbständigung,* Campus, Frankfurt, pp. 11-44

Mayntz, R. (1990) *The Influence of Natural Science Theories on Contemporary Social Science,* MPIFG Discussion Paper, Max-Planck-Institut für Gesellschaftsforschung, Köln

Mayntz, R., Rosewitz, B., Schimank, U., and Stichweh, R. (1988) *Differenzierung und Verselbststständigung.* Campus, Frankfurt

McCarthy, T. (1986) Komplexität und Demokratie - Die Versuchung der Systemtheorie, in A. Honneth and H. Joas (eds.), *Kommunikatives Handeln,* Suhrkamp, Frankfurt, pp. 177-215

McClennen, E.E., (1977) Pragmatic Rationality and Rules *Zeitschrift für Wissenschaftsforschung 11,* 153-186

McCulloch, W.S. and Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys. 5,* 115

Michalewicz, Z. (1994) *Genetic Algorithms + Data Structures = Evolution Programs,* Springer, Berlin

Milgram, S. (1967) The small world problem, *Psychology Today 1,* 61-67

Montana, D.J. (1995) Neural Network Weight Selection Using Genetic Algorithms, in S. Goonatilake and S. Khebbal (eds.), *Intelligent Hybrid Systems,* John Wiley & Sons, London

Moreno, J.L. (1934) Who Shall Survive *Nervous and Mental Disease Monograph 58,* Washington D.C.

Müller, R.A. (1991) *Der (un)teilbare Geist. Modularismus und Holismus in der Kognitionsforschung,* De Gruyter, Berlin

Münch, R. (1988) *Theorie des Handelns,* Suhrkamp, Frankfurt

Needham, J. (1970) *Clercs and Craftsmen in China and the West,* Cambridge University Press, Cambrige

Nowak, M.A. and May, R.M. (1993) The Spatial Dilemma of Evolution, *International Journal of Bifurcation and Chaos 3,* 35-78

Nowak, A. and Lewenstein, M. (1996) Modeling Social Change with Cellular Automata, in R. Hegselmann, U. Mueller and K.G. Troitzsch (eds.), *Modeling and Simulation in the Social Sciences from the Philosophy of Science Point of View,* Kluver Academic Publishers, Dordrecht

Oevermann, U. (1991) Genetischer Strukturalismus und das sozialwissenschaftliche Problem der Erklärung der Entstehung des Neuen, in S. Müller-Dohm (ed.), *Jenseits der Utopie,* Suhrkamp, Frankfurt

Packard, N.H., 1988: Adaptation Towards the Edge of Chaos, in J.A.S. Kelso, A.J. Mandell, and M.F. Shlesinger (eds.), *Dynamic Patterns in Complex Systems,* World Scientific, Singapore, pp. 293-301

Parisi, D., Ceconi, F., and Cerini, A. (1995) Kin-directed altruism and attachment behaviour in an evolving population of neural networks, in N. Gilbert, R. Conte (eds.), *Artificial Societies. The Computer Simulation of Social Life,* UCL Press, London, 238-251

Parsons, T. and Platt, G. (1973) *The American University,* Harvard University Press, Cambridge (Mass.)

Passerini, E. and Bahr, D. (1997) Collective Behaviour Following Disasters: A Cellular Automaton Model, in R.A. Eve, S. Horsfall, and M.E. Lee,*Chaos, Complexity, and Sociology. Myths, Models, and Theories,* Sage, London, pp. 215-228

Pearson, W.W., Steele, N.C., and Albrecht, R.F. (1995) *Artificial Neural Nets and Genetic Algorithms,* Proceedings of the International Conference in Alès, France. Springer, Wien

Peat, F. D. (1988) *Superstrings and the Search for the Theory of Everything,* Contemporary Books, Chicago

Penrose, R. (1989) *The Emperor's New Mind,* Oxford University Press, New York

Piaget, J. (1975) *Die Entwicklung des Erkennens,* Klett, Stuttgart

Pimm,S. (1991) *The Balance of Nature,* University of Chicago Press, Chicago

Pinker, S. (1994) *The Language Instinct,* Morrow, New York

Popper, K.R. (1969) *Die Logik der Forschung,* Mohr, Tübingen

Popper, K.R. and Eccles, J. (1977) *The Self and Its Brain: An Argument for Interactionism,* Springer, Heidelberg

Poundstone, W. (1987) *The Recursive Universe,* Oxford University Press, Oxford

Poundstone, W. (1995) *Das Labyrinth des Denkens*, Rowohlt, Reinbek

Rapoport, A. and Chammah, A.M. (1965) *Prisoner's Dilemma,* University of Michigan Press, Ann Arbor

Rapoport, A. (1986) *General System Theory,* Kent and Cambridge, Tunbridge Wells (Mass.)

Rasmussen, S., Knudsen, C., and Feldberg, R. (1992) Dynamics of Programmable Matter, in C.G. Langton , C. Taylor, J.D. Farmer, and S. Rasmussen (eds.), *Artificial Life II,* Addison Wesley, Reading (Mass.), pp. 211-254

Ray, T.S. (1992) An Approach to the Synthesis of Life, in C.G. Langton , C. Taylor, J. D. Farmer, and S. Rasmussen (eds.), *Artificial Life II,* Addison Wesley, Reading (Mass.), pp. 371-408

Read, D. (1990) The Utility of Mathematical Constructs in Building Archaeological Theory, in A. Voorrips (ed), *Mathematics and Information Science in Archaeology*, Holos, Bonn, pp. 29-60

Rechenberg, I. (1973) *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution,* Frommann-Holzboog, Stuttgart

Ritter, H., Kohonen, T. (1990) Learning "Semantotopic Maps" from Context, *Proceedings of the International Joint Conference on Neural Networks, IJCNN-90, I,* 23-26

Roth, G. (1987) Autopoiese und Kognition. Die Theorie H.R. Maturanas und die Notwendigkeit ihrer Weiterentwicklung, in S. J. Schmidt (ed.), *Der Diskurs des Radikalen Konstruktivismus,* Suhrkamp, Frankfurt, pp. 256-286

Roth, G. (1996) *Das Gehirn und seine Wirklichkeit. Kognitive Neurobiologie und ihre philosophischen Konsequenzen,* Suhrkamp, Frankfurt

Schelling, T.C. (1971) Dynamical Models of Segregation, *Journal of Mathematical Sociology 1,* 143-186

Schmidt, S.J. (ed.) (1987) *Der Diskurs des Radikalen Konstruktivismus,* Suhrkamp Frankfurt

Schnelle, H. (1991) *Die Natur der Sprache,* de Gruyter, Berlin

Schöneburg, E., Hansen, N., and Gawelczyk, A. (1990) *Neuronale Netzwerke,* Markt und Technik, München

Schütz, A. (1971) *Gesammelte Aufsätze Bd. I,* Martinus Nijhoff, Den Haag

Schwartzman, H. (1978) *Transformations. The Anthropology of Children's Play,* Plenum New York

Schweizer, T. (ed.) (1989) *Netzwerkanalyse: Ethnologische Perspektiven,* Reimer, Berlin

Schweizer, T. (1989) Netzwerkanalyse als moderne Strukturanalyse, in T. Schweizer (ed.), *Netzwerkanalyse: Ethnologische Perspektiven,* Reimer, Berlin, pp. 1-32

Shannon, C.E. and Weaver, W. (1963) *The Mathematical Theory of Information,* University of Illinois Press, Urbana

Skvoretz; J. and Fararo, T., 1995: The Evolution of Systems of Social Interaction, *Current Perspectives in Social Theory 15,* 275-299

Spengler, T. (1979) Die Entdeckung der chinesischen Wissenschafts- und Technikgeschichte, in T. Spengler (ed.), *Joseph Needham, Wissenschaftlicher Universalismus,* Suhrkamp, Frankfurt, pp. 7-52

Sperber, D. (1982) *On Anthropological Knowledge,* Cambridge University Press, Cambridge (Mass.)

Stadler, M. and Kruse, P. (1992) Zur Emergenz psychischer Qualitäten. Das psycho-physische Problem im Lichte der Selbstorganisationstheorie, in W. Krohn, G. Küppers (eds.), *Emergenz: Die Entstehung von Ordnung, Organisation und Bedeutung,* Suhrkamp, Frankfurt, pp. 134-160

Stanley, E.A., Ashlock, D., Tesfatsion, L. (1994) Iterated Prisoner's Dilemma and Refusal of Partners, in C.G. Langton (ed.), *Artificial Life III,* Addison Wesley, Reading (Mass.), pp. 131-175

Stegmüller, W. (1957) *Das Wahrheitsproblem und die Idee der Semantik,* Springer,Berlin

Stoica, C. (1997) *Die Vernetzung sozialer Einheiten. Hybride interaktive neuronale Netzwerke in den Kommunikations- und Sozialwissenschaften,* Dissertation Essen

Sylvan, D. and Glassner, B. (1987) *A Rationalist Methodology for the Social Sciences,* Basil Blackwell, London

Tarski, A. (1956) *Logics, Semantics, Metamathematics,* Oxford University Press, Oxford

Tipler, F. J. (1994) *The Physics of Immortality,* Double Day, New York

Toffoli, T. (1997) How much can you get for how little? A conceptual introduction to cellular automata, *Interjournal,* New England Complex Systems Institute, Boston

Troitzsch, K.G. (1990) *Modellbildung und Simulation in den Sozialwissenschaften,* Westdeutscher Verlag, Opladen

Troitzsch, K.G. (1995) Chaotic Behaviour in Social Systems, in R. Hegselmann and H.O. Peitgen (eds.), *Modelle sozialer Dynamiken - Ordnung, Chaos, Komplexität,* Wien, pp. 162-186

Vichniac, G. (1984) Simulating Physics with Cellular Automata *Physica **10 D,*** 96-116

Waldrop, M. M. (1992) *Complexity. The Emerging Science at the Edge of Order and Chaos,* Simon & Schuster, New York

Watzlawick, P., Beavin, J.H., and Jackson, D.D. (1985) *Menschliche Kommunikation,* Hans Huber, Bern

Weber, M. (1981/1920) *Die protestantische Ethik I,* Mohr, Tübingen

Weber, M. (1982) R. Stammlers Überwindung der materialistischen Geschichtsauffassung, in J. Winckler (ed.), *Gesammelte Aufsätze zur Wissenschaftslehre von Max Weber,* Mohr, Tübingen, pp. 291-359

Weingarten, E. and Sack, F. (1976) Ethnomethodologie. Die methodische Konstruktion der Realität, in E. Weingarten, F. Sack, and J. Schenkein (eds.), *Ethnomethodologie. Beiträge zu einer Soziologie des Alltagshandelns,* Suhrkamp, Frankfurt, pp. 7-26

Willke, H. (1995) *Systemtheorie III: Steuerungstheorie,* Gustav Fischer, Stuttgart, Jena

Wilson, E.O. (1975) *Sociobiology. The New Synthesis,* Harvard University Press, Cambridge (MA)

Wolfram, S. (ed.) (1986) *Theory and Applications of Cellular Automata,* World Scientific, Singapore

Wuensche, A. and Lesser, M. (1992) *The Global Dynamics of Cellular Automata: Attractions Fields of One-Dimensional Cellular Automata,* Addison Wesley, Reading (MA)

Wuensche, A. (1994) The Ghost in the Machine: Basins of Attraction of Random Boolean Networks, in C.G. Langton (ed.), *Artificial Life III,* Addison Wesley, Reading (Mass.), pp. 465-501

Yoffee, N. (1994) Memorandum to Murray Gell-Mann Concerning: The Complications of Complexity in the Prehistoric Southwest, in G.J. Gumermann and M. Gell-Mann (eds.), *Understanding Complexity in the Prehistoric Southwest,* Addison Wesley, Reading (Mass.), pp 341-358

Zilsel, E. (1976) Die sozialen Ursprünge neuzeitlicher Wissenschaft. Suhrkamp,Frankfurt

INDEX

A

action 1, 7, 34, 56, 76p,
actor 24, 76, 198, 250p
adaptation 9, 17, 30, 32pp, 74p, 106, 114pp, 123p, 126, 148p, 161p, 251, 263
aggressive strategy 118, 192, 199, 206
AGIL-scheme 49
algorithmic complexity 58pp, 63, 87p, 267
anticipation 46, 76, 154, 201, 204, 206
Apel, K.O. 6, 18, 135, 252, 257
autocatalysis 113, 235p
Axelrod, R. 181, 183, 186pp, 191pp, 200p, 204, 206

B

Bateson, G. 17, 23, 34, 36p, 39p, 47, 68p,116, 123, 128, 158, 175, 178, 206, 231, 243p, 269
Boolean Networks (BNs) 96pp, 118, 124, 148, 158, 212, 219
Bourdieu, P. 174, 215, 248pp
bourgeois society 186
brain 10p, 18, 33, 41pp,56, 66, 126, 220, 237, 262, 264
Brown, D.E. 257, 258, 259, 261p

C

Cellular Automata (CAs)
- hybrid 149, 152, 200pp
- deterministic 92, 94
Cantor, G. 51, 53
Chaitin, G.J. 53, 58
chaos 5, 7, 25pp, 64, 72, 86p, 90, 95, 104, 231p
chaotic system 12, 27p, 60pp, 90p, 154, 232
Chomsky, N. 15, 16p, 81, 124, 126, 140, 160, 164, 173, 248, 257p, 260, 262
Churchland, P. 7, 10p, 15, 20, 264, 269
class society 31, 129
complexity
- algorithmic 58pp, 63, 87p, 267
- effective 59, 60pp, 87, 267
- class 113, 151, 177, 189
- degree of complexity 60pp
- space 61p
computability 50pp, 68, 87, 162
contractive function 111p
control parameter 81, 83, 86, 89, 93pp, 100pp, 140pp, 145pp,150p, 167, 189p, 194, 211, 218, 254, 260, 269
convergence 111pp, 212
cooperative strategy 118, 120, 187, 199
crossover 107pp, 113pp, 133, 139, 141, 143p, 156, 177, 199p, 234

D

Darwin, C. 115, 256
Dawkins, R. 110, 116, 169, 255
degree of complexity 60pp
democracy 99
deterministic CAs 92, 94
deterministic system 28, 87, 91p, 94
dialectics 1, 66, 271

dimension 55, 57, 84p, 103, 131, 159, 162, 180, 223, 225pp, 231, 233
Dreyfus, H.L. 161, 205, 249, 253
dynamics 25, 79, 81p, 86, 90, 92pp, 113, 122pp, 127p, 133p, 136p, 139pp, 145, 147, 150, 152pp

E

ecology 95
ecosystem 114
Edelman, G.M. 11, 33, 42, 220, 237
Eder, K. 69, 162, 225, 235
effective complexity 59, 60pp, 87, 267
emergence 9, 50, 56, 63pp, 79, 147, 179, 214, 226, 236
entropy 57p
environmental vector 166, 176p, 202, 231
evaluation function 33, 108p, 111, 113, 130pp, 160, 166, 168, 174, 229

F

Fararo, T. 3, 161, 194
feudal society 167
Freeman, L. 102, 214p
function
- contractive 111p
- evaluation function 33, 108p, 111, 113, 130pp, 160, 166, 168, 174, 229
- transition 25p, 85, 89, 90, 93, 97, 108pp
functional differentiation 225p, 230p
functionalism 73, 74

G

Genetic Algorithms (GAs) 106pp, 116, 118, 123, 140, 142, 148, 156, 191, 193
Geertz, C. 19
Gell-Mann, M. 8, 10pp, 23, 59, 63, 75, 87, 165
geometry 92, 102, 106, 180p, 210pp, 219p, 227p
Giddens, A. 248
Gilbert, N. 24
Goonatilake, S. 39, 124, 128
Grossberg, S. 47p

H

Habermas, J. 1, 5, 69, 104, 139, 164, 182, 225p, 231, 242, 246, 252, 254, 257, 259p, 262p, 264pp
Hawking, S.W. 224
Hebb, D.O. 39, 43, 151
Hegselmann, R. 81, 162, 221
Hillis, D. 108, 112
history 42, 48, 50p, 67, 74pp, 115, 126, 137, 154, 162, 164, 167p, 181, 192, 195, 197p, 202, 206, 208, 225, 239, 257
Hobbes, T. 186, 194, 197, 208, 261
Hofstadter, D.R. 18, 35, 54, 69, 122, 139, 140, 151, 242
Holland, J.R. 48, 83, 107p, 110, 127, 160, 169p
hybrid CAs 149, 152, 200pp
hybrid system 31, 76, 113, 115, 121, 122pp, 132p, 135p, 137, 140pp, 150p, 153pp, 162, 164p, 170, 175, 186, 201, 237, 243p, 251, 268pp

I

institution 138, 174, 259
interaction 29pp, 82, 97, 109, 112, 118, 122p, 155, 157

K

Kant, E. 72, 80, 181, 224, 228, 241
Kantian stance 72, 80, 139, 170, 244
Kauffman, S.A. 70pp, 80, 89p, 96pp, 103pp, 113, 116, 139, 144, 170, 184
Khebbal, S. 39, 124, 128

Koza, J.R. 39, 80, 107, 108, 124
Krohn, W. 29p, 64, 65, 162
Küppers, G. 29p, 64, 65, 162

L

Langton, C.G. 7pp, 13, 30, 58, 82, 86pp, 98, 231
Lansing, J.S. 184, 187
learning 8, 18, 22, 36pp, 66p, 76, 106, 121pp, 127, 138, 148, 153, 161p, 170, 175, 178, 206, 243, 247, 262, 264
learning rule 36, 38pp, 66, 123, 170, 247
linear system 25, 27, 247
linguistic universals 104, 124p, 257p, 262
logical level 23, 33p, 36, 54p, 69, 122, 126p, 141, 146p, 151, 186
logical rule 153, 247
Luckmann, T. 66, 134, 246, 257
Luhmann, N. 1p, 4pp, 15p, 18, 24, 53, 55, 66, 74, 130, 136, 157pp, 163, 177, 184, 186, 225pp, 248, 254

M

Machiavelli, N. 271
Marx, K. 1p, 5, 61, 86, 135pp, 164, 184, 186, 232pp, 263, 265p
mathematical sociology 3, 161, 181, 241, 245p
Maturana, H.R. 29, 33, 138
May, R.M. 114, 187pp, 194pp, 208p, 252, 261
Maynard Smith, J. 72, 80, 182p, 190, 192p, 204, 208, 245
Mayntz, R. 6, 161, 225
memory 41pp, 71, 151pp, 243p
meta-attractor 146
meta-level 54pp, 159
meta-parameter 142pp, 151, 255, 261, 263, 267
meta-rule 39
metric 210pp, 214, 219p, 226p
Michalewicz, Z. 106p,110, 112, 168
Milgram, S. 215, 221
mind 3, 10p,16pp, 33p, 41, 51, 53pp, 62, 66, 82, 91, 117, 121, 122, 125, 130, 132, 134, 136, 138p, 150, 154, 179, 206, 210, 243p, 255
modern society 233p, 265
Moreno, J.L. 217
mutation 8, 13, 75, 107pp, 130, 132, 135, 138, 141, 143, 146, 149, 156, 177, 191, 197, 199, 208, 221, 234, 255

N

neighborhood 83pp, 93, 101, 128p, 187pp, 211p, 215p, 221p, 252
network analysis 102, 213, 215, 218, 220
neural nets 11, 39, 43p, 66, 70, 83, 96, 107, 123, 126, 128, 129, 144, 147, 170, 175, 220
non-linear system 25pp, 83
Nowak, M.A. 81, 187pp, 194pp, 208p, 252, 261

O

optimization 31, 39, 106pp, 116, 127, 130, 132, 135, 142, 147pp, 168, 170, 177pp, 204, 206p, 236, 270

P

Packard, N.H. 86, 87, 90, 104, 231
parameter
- control 81, 83, 86, 89, 93pp, 100pp, 140pp, 145pp,150p, 167, 189p, 194, 211, 218, 254, 260, 269
- meta 142pp, 151, 255, 261, 263, 267
Parsons, T. 5, 49p, 174, 246

pay-off matrix 117, 182p 185, 188, 190, 194pp, 208p, 252
phase space 211
Piaget, J. 17, 32, 74, 164
Pinker, S. 104, 124p, 258p, 262
Popper, K.R. 63, 64, 172, 255
potential complexity 63
predator-prey system 21p, 26, 80, 95, 209

R

Rapoport, A. 117p, 181, 186, 194p, 201, 218
Read, D. 246
role 8, 12p, 18, 50, 53, 70, 81, 90, 116, 136, 157, 184p, 225pp
Roth, G. 18, 33, 45, 66, 114, 220
rule
- learning 36, 38pp, 66, 123, 170, 247
- logical 153, 247
- meta-rule 39
- social 50, 95, 134p, 137, 157, 163, 174, 234, 237, 241, 248p, 252pp, 259pp
- totalistic 85, 128

S

Schütz, A. 75pp, 246
segmentary differentiation 162, 225p, 230p, 244
Sejnowski, T. 7, 10p, 15, 20
selection 8, 13, 29pp, 107, 112, 113, 118, 120, 130, 156, 161, 192, 198p, 201, 202, 255
self-enlargement 122, 129, 236pp
self-organization 5, 7p, 16p, 22, 28pp, 70, 74p, 90, 96, 101, 104, 106, 115p, 122, 124pp, 158, 161, 242
self-referentiality 18, 264, 269
set theory 84, 158, 211
Simmel, G. 220, 229
simple attractor 43, 46, 87, 91, 98, 141, 142, 153, 231, 255, 260, 269
social differentiation 70, 130, 132, 142, 166, 168, 174, 184, 204p, 227
social level 19, 70, 161, 186, 229p
social rule 1, 50, 95, 134p, 137, 157, 163, 174, 234, 237, 241, 248p, 252pp, 259pp
social universals 241, 243, 257, 259pp
society
- bourgeois society 186
- class society 31, 129
- feudal society 167
- modern 233p, 265
state space 15, 25pp, 62, 85, 92, 106, 112, 211, 224
stochastic CAs 92, 93
stochastic system 93, 94, 109, 128, 140, 160, 170, 173, 191, 196, 231, 251
strange attractor 27, 86p, 153p, 231
strange loop 69, 242
strategy
- aggressive strategy 118, 192, 199, 206
- cooperative strategy 118, 120, 187, 199
stratificatory differentiation 162, 225, 23op
subsystem 19, 43p, 174, 179, 234
system
- chaotic 12, 27p, 60pp, 90p, 154, 232
- deterministic 28, 87, 91p, 94, 160, 170
- hybrid 31, 76, 113, 115, 121, 122pp, 132p, 135p, 138, 140pp, 150p, 153pp, 162, 164p, 170, 175, 186, 201, 237, 243p, 251, 268pp
- linear 25, 27, 247

system function 13, 26p, 85, 97, 101, 130, 134, 154, 160, 166, 184
system value 130, 133, 141p, 160, 166, 171, 176, 199pp

T

theory of games 36, 72, 93, 117, 118, 181pp, 193, 199, 204, 209, 221, 250
topology 39, 99, 101, 110, 210pp, 215, 219p, 227
totalistic rule 85, 128
trajectory 26pp, 58, 62, 87, 92, 112p, 134, 153pp, 167p, 211, 254, 270
transition function 25p, 85, 89, 90, 93, 97, 108pp
Turing, A. 14, 50pp, 56, 81p

V

Varela, F.J. 33, 56, 126
variability 33, 44, 47p, 62, 114, 126, 137, 139, 140, 142, 146, 147, 149pp, 154, 157, 170, 200, 202, 204pp, 232, 234, 238p, 244, 246p, 252, 254, 259, 263, 266pp, 271

W

Weber, M. 1, 5, 186, 226, 246, 249pp, 265p
Wolfram, S. 82pp, 97pp, 112, 115, 134, 136p, 140p, 145, 154, 190, 221
Wolfram class 61, 83, 86p, 90pp, 97pp, 112, 115, 134, 136p, 140p, 145, 154, 189, 221
Wuensche, A. 46, 48, 82, 91p,101, 105, 153, 211

THEORY AND DECISION LIBRARY

SERIES A: PHILOSOPHY AND METHODOLOGY OF THE SOCIAL SCIENCES

Editors: W. Leinfellner (*Vienna*) and G. Eberlein (*Munich*)

1. G. Zecha and P. Weingartner (eds.): *Conscience: An Interdisciplinary View*. Salzburg Colloquium on Ethics in the Sciences and Humanities (1984). 1987 ISBN 90-277-2452-0
2. R.W. Scholz: *Cognitive Strategies in Stochastic Thinking*. 1987 ISBN 90-277-2454-7
3. H. Nurmi: *Comparing Voting Systems*. 1987 ISBN 90-277-2600-0
4. M. Schmid and F.M. Wuketits (eds.): *Evolutionary Theory in Social Science*. 1987 ISBN 90-277-2612-4
5. C.L. Sheng: *A New Approach to Utilitarianism*. A Unified Utilitarian Theory and Its Application to Distributive Justice. 1991 ISBN 0-7923-1301-1
6. F. Forman: *The Metaphysics of Liberty*. 1989 ISBN 0-7923-0080-7
7. G. Bernard: *Principia Economica*. 1989 ISBN 0-7923-0186-2
8. J.W. Sutherland: *Towards a Strategic Management and Decision Technology*. Modern Approaches to Organizational Planning and Positioning. 1989 ISBN 0-7923-0245-1
9. C. Vlek and G. Cvetkovich (eds.): *Social Decision Methodology for Technological Projects*. 1989 ISBN 0-7923-0371-7
10. P. Hoyningen-Huene and F.M. Wuketits (eds.): *Reductionism and Systems Theory in the Life Sciences*. Some Problems and Perspectives. 1989 ISBN 0-7923-0375-X
11. K.G. Grunert and F. Ölander (eds.): *Understanding Economic Behaviour*. 1989 ISBN 0-7923-0482-9
12. G. Antonides: *The Lifetime of a Durable Good*. An Economic Psychological Approach. 1990 ISBN 0-7923-0574-4
13. G.M. von Furstenberg (ed.): *Acting under Uncertainty*. Multidisciplinary Conceptions. 1990 ISBN 0-7923-9063-6
14. R.L. Dukes: *Worlds Apart*. Collective Action in Simulated Agrarian and Industrial Societies. 1990 ISBN 0-7923-0620-1
15. S.O. Funtowicz and J.R. Ravetz: *Uncertainty and Quality in Science for Policy*. 1990 ISBN 0-7923-0799-2
16. J. Götschl (ed.): *Erwin Schrödinger's World View*. The Dynamics of Knowledge and Reality. 1992 ISBN 0-7923-1694-0
17. R. von Schomberg (ed.): *Science, Politics and Morality*. Scientific Uncertainty and Decision Making. 1993 ISBN 0-7923-1997-4
18. J. Baron: *Morality and Rational Choice*. 1993 ISBN 0-7923-2276-2
19. M. Allais and O. Hagen (eds.): *Cardinalism*. A Fundamental Approach. 1994 ISBN 0-7923-2398-X
20. E. Carlson: *Consequentialism Reconsidered*. 1995 ISBN 0-7923-3547-3
21. J. Götschl (ed.): *Revolutionary Changes in Understanding Man and Society*. 1995 ISBN 0-7923-3627-5
22. A. Camacho: *Division of Labor, Variability, Coordination, and the Theory of Firms and Markets*. 1996 ISBN 0-7923-3832-4
23. R. Hegselmann, U. Mueller and K.G. Troitzsch (eds.): *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*. 1996 ISBN 0-7923-4125-2

24. J. Nida-Rümelin: *Economic Rationality and Practical Reason.* 1997 ISBN 0-7923-4493-6
25. G. Barbiroli: *The Dynamics of Technology.* A Methodological Framework for Techno-Economic Analyses. 1997 ISBN 0-7923-4756-0
26. O. Weinberger: *Alternative Action Theory.* Simultaneously a Critique of Georg Henrik von Wright's Practical Philosophy. 1998 ISBN 0-7923-5184-3
27. A. Carsetti (ed.): *Functional Models of Cognition.* Self-Organizing Dynamics and Semantic Structures in Cognitive Systems. 2000 ISBN 0-7923-6072-9
28. J. Nida-Rümelin and W. Spohn (eds.): *Rationality, Rules, and Structure.* 2000 ISBN 0-7923-6326-4
29. J. Klüver: *The Dynamics and Evolution of Social Systems.* New Foundations of a Mathematical Sociology. 2000 ISBN 0-7923-6443-0